

FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO



Análise e Visualização de Interacções Online em Redes Sociais

Luis Carlos Rijo Gaspar

Relatório de Dissertação

Mestrado Integrado em Engenharia Informática e Computação

Orientador: Álvaro Reis Figueira (Professor Doutor)

Março de 2010

Análise e Visualização de Interações Online em Redes Sociais

Luis Carlos Rijo Gaspar

Relatório de Dissertação

Mestrado Integrado em Engenharia Informática e Computação

Aprovado em provas públicas pelo Júri:

Presidente: Ana Paula Rocha (Professora Auxiliar da FEUP)

Arguente: Carlos Vaz de Carvalho (Professor Adjunto do ISEP)

Vogal: Álvaro Reis Figueira (Professor Assistente da FCUP)

24 de Março de 2010

Resumo

Neste trabalho implementámos uma aplicação que permite analisar interacções em fóruns de discussão *online*. Esta análise é efectuada em duas vertentes: a análise visual e a análise numérica. A análise visual é feita através da visualização de um grafo que representa as interacções entre actores (redes de 1-Modo) ou as interacções entre actores e discussões de um fórum (redes de 2-Modos). Este tipo de análise permite “olhar” para as interacções sob uma perspectiva qualitativa. A análise numérica é realizada tendo por base vários indicadores numéricos de análise, nomeadamente, Grau de Centralidade, Cliques, Densidade e Índice de Centralização. O primeiro enquadra-se numa análise local, descrevendo as interacções numa perspectiva individual. O segundo diz respeito a uma análise grupal, analisando possíveis grupos que existam dentro de um determinado fórum. Os dois últimos indicadores são referentes a uma análise global, servindo para analisar as interacções entre actores como um conjunto, quer seja com todos os actores ou só com aqueles que realmente participam no fórum.

Este trabalho baseia-se na teoria da Análise de Redes Sociais e contribui para que a análise a fóruns de discussão *online* possa ser facilitada, através da utilização da aplicação desenvolvida. Esta aplicação pode ser utilizada num contexto de ensino, mais propriamente, num *Learning Management System*, como é o sistema Moodle.

Abstract

In this work we create an application for analyzing interactions in *online* discussion forums. This analysis is conducted in two parts: a visual analysis and numerical analysis. The visual analysis is done by using a graph representing the interactions between users in forums (networks of 1-Mode) or the interactions between users and a forum threads (2-Mode). This type of analysis allows a qualitative look to the interactions. The numerical analysis is performed based on several indicators, in particular, Centrality Degree, Cliques, Density and Centralization Index. The first indicator is part of a local analysis, describing the interactions in an individual perspective. The second, concerns to a group analysis, identifying possible groups that exist within a particular forum. The last two indicators are related to a global analysis, describing the interactions between users as a whole, whether it be with all users or only those who actually participate in the forum.

This work is based on the theory of Social Network Analysis and contributes to the analysis of *online* discussion forums. It can be facilitated through the use of the application developed. This application can be used in a context of education, more specifically in Moodle, which is a Learning Management System.

Agradecimentos

O desenvolvimento do presente trabalho não teria sido possível sem a colaboração de algumas pessoas.

Ao Professor Doutor Álvaro Reis Figueira, pela orientação e pelo empenho que mostrou para que o trabalho fosse bem sucedido.

Ao Professor Doutor João Correia Lopes, por me ter disponibilizado a wiki para alojar a página da dissertação.

Ao Investigador Nuno Miguel Barbosa, pela ajuda que me deu na instalação das ferramentas ligadas à utilização do Moodle.

Aos meus pais, Maria da Conceição e Luis Manuel Gaspar, pelo apoio e incentivo ao longo deste trabalho.

Aos meus amigos, Gustavo Martins, Hector Dantas, Hugo Gomes e Vitor Pereira, pela amizade, pelos almoços animados e por todos os momentos passados.

Ao meu amigo Sergio Ferreira e ao meu primo Francisco Gaspar, pela ajuda na correcção ortográfica deste trabalho.

A todos os meus amigos e companheiros de “futebolada” aos fins de semana, por estarem sempre disponíveis para esses momentos tão importantes de descontração.

A todos aqueles que, ao longo deste período, contribuíram com a sua opinião, crítica ou sugestão e que, de certa forma, influenciaram a sua realização.

Índice

1	Introdução.....	1
1.1	Contexto/Enquadramento.....	1
1.2	Motivação e Objectivos.....	2
1.3	Estrutura da Dissertação.....	3
2	Revisão Bibliográfica.....	4
2.1	Fóruns de Discussão.....	4
2.2	Análise de Redes Sociais.....	5
2.2.1	Conceito de ARS.....	5
2.2.2	Redes de 1-Modo.....	5
2.2.3	Redes de 2-Modos.....	8
2.2.4	Níveis de Análise.....	9
2.3	Tecnologias adoptadas e justificação.....	10
2.3.1	Linguagem de desenvolvimento da aplicação.....	10
2.3.2	Sistema de Gestão de Bases de Dados.....	10
2.3.3	Linguagem de desenvolvimento web.....	11
2.4	Tecnologias existentes para ARS.....	14
2.4.1	UCINet.....	14
2.4.2	NetDraw.....	15
2.4.3	Graphviz.....	15
2.4.4	Pajek.....	15
2.4.5	Comparações.....	15
3	Representação de Grafos e Indicadores Numéricos de Análise.....	16
3.1	Indicadores numéricos de análise.....	16
3.1.1	Análise Local.....	17
3.1.2	Análise Grupal.....	20
3.1.3	Análise Global.....	20
3.2	Redes de 1-Modo.....	25
3.3	Redes de 2-Modos.....	26
3.4	Conceitos importantes do algoritmo de desenho do grafo.....	27
4	O Protótipo.....	28
4.1	Arquitectura da aplicação.....	29

4.2	Script de recolha de dados.....	30
4.2.1	Descrição do algoritmo utilizado.....	30
4.2.2	Demonstração da execução do script de recolha de interacções.....	31
4.3	Implementação do baixo nível.....	32
4.3.1	Leitura e armazenamento do ficheiro de interacções.....	32
4.3.2	Cálculo do diâmetro dos nós do grafo.....	33
4.3.3	Representação das interacções no grafo.....	35
4.3.4	Definição da espessura das interacções no grafo.....	36
4.3.5	Detecção dos ficheiros de fóruns no directório raiz da aplicação.....	39
4.3.6	Movimentação dos nós no grafo.....	40
4.3.7	Algoritmo de desenho do grafo.....	40
4.3.8	Demonstração da execução do algoritmo.....	46
4.4	Visita guiada à interface.....	47
4.4.1	Representação do grafo.....	48
4.4.2	Métricas Locais.....	49
4.4.3	Métricas Globais.....	50
4.4.4	Informação sobre actores.....	51
4.4.5	Escolher fórum.....	51
4.4.6	Escolher tipo de rede.....	52
4.4.7	Opções disponíveis.....	52
4.4.8	Ajuda.....	53
4.5	Applet.....	53
5	Conclusões e Trabalho Futuro.....	54
5.1	Resumo do trabalho.....	54
5.2	Contribuições.....	56
5.3	Fragilidades detectadas.....	57
5.4	Trabalho Futuro.....	57
	Referências.....	58

Lista de Figuras

Figura 3:1: Exemplo de um grafo representado na interface da aplicação.....	18
Figura 3:2: Rede em Estrela (Índice de Centralização igual a 100%).....	23
Figura 3:3: Rede sem um actor central (Índice de Centralização igual a 0%).....	23
Figura 3:4: Grafo sob a forma de rede de 1-Modo.....	25
Figura 4:1: Arquitectura da aplicação.....	29
Figura 4:2: Encadeamento de mensagens num fórum de discussão.....	31
Figura 4:3: Ficheiro de output.....	31
Figura 4:4: Representação de uma interacção entre o mesmo actor.....	36
Figura 4:5: Interacções entre os actores C e G.....	38
Figura 4:6: Interacções entre os actores A e D.....	38
Figura 4:7: Interacção entre os actores A e C.....	39
Figura 4:8: Representação de um grafo, utilizando o algoritmo apresentado.	46
Figura 4:9: Interface da aplicação desenvolvida.....	47
Figura 4:10: Frame de representação do grafo.....	48
Figura 4:11: Métricas locais.....	49
Figura 4:12: Métricas globais.....	50
Figura 4:13: Dados dos actores.....	51
Figura 4:14: ComboBox para escolha do fórum.....	51
Figura 4:15: ComboBox para escolha da rede.....	52

Lista de Tabelas

Tabela 2.1: Disposição de mensagens de um fórum de discussão, em níveis hierárquicos.....	6
Tabela 2.2: Matriz binária das interações entre actores.....	7
Tabela 2.3: Matriz binária das interações entre actores.....	8
Tabela 2.4: Matriz de representação das participações em discussões.....	9
Tabela 3.1: Grau de Centralidade, tendo em conta a existência de ligações.....	19
Tabela 3.2: Grau de Centralidade, considerando o peso das ligações.....	19
Tabela 3.3: Densidade do grafo para participantes activos ou rede completa.....	22
Tabela 3.4: Índice de Centralização, para a rede completa e participantes activos.....	24

Abreviaturas e Símbolos

ARS	Análise de Redes Sociais
GC	Grau de Centralidade
IC	Índice de Centralização
PHP	Hypertext Preprocessor
ASP	ActiveServer Pages
JSP	JavaServer Pages
LMS	Learning Management System
SGBD	Sistema de Gestão de Bases de Dados
HTML	HyperText Markup Language

1 Introdução

Neste capítulo explicamos o contexto/enquadramento da proposta que serve de base a esta tese de Mestrado. Neste ponto são apresentadas informações sobre a proposta, apresentando-se ainda a área específica relacionada com a proposta e uma breve descrição desta. Seguidamente, são apresentadas as justificações para a motivação da escolha da proposta referida, bem como os objectivos que essa proposta possui em relação à sua elaboração. Finalmente, é apresentada a estrutura da dissertação, sendo apresentados e descritos em resumo os capítulos que a compõem.

1.1 Contexto/Enquadramento

O presente trabalho visa a implementação de uma aplicação, com vista a permitir a análise das interacções em fóruns de discussão *online* embebida num LMS e está ligada à área de ARS. Esta proposta tem por base uma tese de Mestrado, Laranjeiro (2008), intitulada "Contributos para a Análise e Caracterização de Interacções em Fóruns de Discussão Online" da autoria de Joanne Brás Laranjeiro e datada de Janeiro de 2008. Esta tese teve como orientadores o professor Álvaro Figueira e a professora Luísa Aires. Nela são descritos um conjunto de conceitos de ARS e que serviram de base à implementação da aplicação pretendida, como sendo as métricas de análise das interacções, os tipos de representações de interacções em fóruns de discussão *online* e as diversas conclusões a tirar, associadas à representação destas interacções.

A área de ARS está, tal como o nome indica, relacionada com a análise a redes sociais. Neste caso concreto, esta aplica-se à análise de fóruns de discussão *online*. O conceito de ARS consiste em vários procedimentos matemáticos, ligados à selecção de uma amostra e à posterior recolha e registo dos dados relativos a essa amostra, com o objectivo de caracterizá-la, tendo em conta diferentes parâmetros de análise.

1.2 Motivação e Objectivos

A área do presente trabalho tem enormes potencialidades no que toca à análise de interacções entre actores, num ambiente *online*, bem como a todas as formas de avaliar essas interacções, podendo, assim, chegar-se a conclusões interessantes ligadas às relações entre pessoas num contexto oposto aos típicos encontros face-a-face.

Para além disso, este trabalho poderá ser adaptado a outros âmbitos de estudo de interacções *online*, como sendo a análise de tendências de actores, ligado ao marketing, e ainda na detecção de padrões de comunicação e identificação dos respectivos actores, ligada à “luta contra o terrorismo”.

Finalmente, este trabalho poderá ser utilizado como base para outros desenvolvimentos, nomeadamente, associação de mecanismos de análise do conteúdo das mensagens trocadas entre actores, no contexto *online*.

Cremos que as conclusões que a aplicação desenvolvida retirará, poderão permitir a obtenção de informação muito importante acerca das personalidades de indivíduos, sendo que nem sequer é necessário conhecê-los pessoalmente para os descrever nalguns pontos da sua personalidade com um grau de certeza elevado, utilizando, para isso, métricas e representação visual de todas as interacções em que estão envolvidos. No contexto do ensino *online*, as informações obtidas pelos professores sobre as interacções dos seus alunos no fórum de uma determinada disciplina, através da utilização desta aplicação, servem para perceber:

- Quais são os alunos mais participativos ou alunos isolados, ou seja, os alunos que participam activamente no fórum, através da colocação de dúvidas ou do esclarecimento das dúvidas dos colegas, ou os alunos que participam pouco ou nada, podendo estes ter dificuldades e não as pretenderem dar a conhecer ou então podem não estar interessados em ajudar os restantes colegas;
- Qual é a “popularidade” de um aluno em relação aos colegas, ou seja, verificar se as mensagens se encaminham em grande número para um único aluno, concluindo-se que este um aluno é “popular” para os colegas;
- Qual é a “influência” de um aluno nos colegas, ou seja, verificar se um determinado aluno envia o maior número de mensagens, podendo dizer-se que é o mais “influyente” na formação de conhecimento dos colegas;
- Se existem grupos de alunos, onde a informação circula com mais intensidade;
- Se a participação no fórum é elevada ou reduzida, ou seja, se o número de interacções entre alunos está próximo do número máximo de interacções possível ou se, pelo contrario, existem poucos alunos a participar e estes interagem com poucos colegas.

Introdução

Os professores, tendo chegado a estas conclusões, podem adoptar medidas para corrigir eventuais problemas com os alunos, nomeadamente, averiguar a pouca participação de um aluno e/ou a baixa participação geral no fórum da sua disciplina.

Quanto aos objectivos da proposta, o primeiro deles é apresentar uma aplicação capaz de permitir a análise de interacções em fóruns de discussão *online*. Esta aplicação tem uma interface onde é possível visualizar a representação de um determinado fórum de discussão e, juntamente com esta representação, são também mostrados diversos indicadores numéricos de análise, visando a complementaridade da informação obtida aquando da visualização da representação de um fórum.

O segundo objectivo é, posteriormente à implementação do primeiro, criar um módulo no Moodle por forma a que a interface implementada seja embebida nesse LMS e assim, permitir aos docentes das várias disciplinas contidas no Moodle a possibilidade de poder analisar o conjunto de contribuições dos seus alunos no âmbito de uma determinada disciplina.

1.3 Estrutura da Dissertação

O presente trabalho foi organizado tendo por base os seguintes capítulos:

- Capítulo 1 – Introdução;
- Capítulo 2 – Revisão Bibliográfica;
- Capítulo 3 – Representação de Grafos e Indicadores Numéricos de Análise;
- Capítulo 4 – O Protótipo;
- Capítulo 5 – Conclusões e Trabalho Futuro;

No capítulo 2 são apresentados os conceitos teóricos que servem de base a este trabalho, bem como as justificações para a tomada de decisão em relação às tecnologias de desenvolvimento associadas, e ainda a justificação para a não utilização de tecnologias já existentes, no âmbito da ARS. No que se refere ao capítulo 3, constam nele a apresentação e descrição das formas de representação das interacções em fóruns de discussão *online* e dos indicadores numéricos de análise associados a essa representação, que visam a sua complementaridade. Em relação ao capítulo 4, este conta com a descrição de tudo o que diz respeito à implementação do trabalho, nomeadamente, o *script* de recolha de dados, a implementação de baixo nível, onde se destaca o algoritmo de desenho automático do grafo, a interface da aplicação desenvolvida e todas as suas funcionalidades e, finalmente, a *applet*. Por último, o capítulo 5 apresenta as conclusões do trabalho elaborado, relativamente aos resultados obtidos e futuros desenvolvimentos.

2 Revisão Bibliográfica

Neste capítulo é descrito o estado da arte do projecto em que se baseia esta tese. Aqui são descritos alguns conceitos que servem de base à fundamentação do projecto, nomeadamente, os conceitos de fóruns de discussão *online* e análise de redes sociais. De seguida são também apresentadas as tecnologias definidas para a implementação da aplicação que permite a análise das interacções entre actores em fóruns de discussão *online*, sendo fundamentadas as opções tomadas. Finalmente, são descritas as tecnologias mais utilizadas no âmbito da ARS, sendo apresentadas as comparações entre elas e o porquê de não se enquadrarem no objectivo deste trabalho.

2.1 Fóruns de Discussão

Os ambientes de e-learning, suportados nos conceitos de interacção e colaboração, não funcionam apenas como simples plataformas de transferência de conteúdos, possibilitando a comunicação entre indivíduos, com vista a acompanhar o seu processo de ensino-aprendizagem.

O desenvolvimento tecnológico é uma constante nos dias que correm, criando inúmeras formas de comunicação entre pessoas. Isto leva a que estas ferramentas permitam uma comunicação em duas variantes: tempo real e tempo diferido. Estas duas variantes são adoptadas, tendo sempre em conta o contexto em que se vai inserir e os objectivos pretendidos para esse contexto. Conclui-se que, para a criação de um ambiente "rico em aprendizagem", é necessária não só a disponibilização de conteúdos para suportar essa aprendizagem, mas também a possibilidade de aprofundar as relações interpessoais, por forma a alicerçar todos os conhecimentos individuais e colectivos.

2.2 Análise de Redes Sociais

A ARS tem-se expandido por vários campos como as ciências sociais, económicas, políticas, saúde, comunicação, redes computacionais/redes de informação, entre outras. Garton *et al.* (1997) [GHW97] referem que a ARS pode aplicar-se a áreas tão diversas como relações familiares, contextos militares e empresariais e combate ao terrorismo. Garton *et al.* (1997) [GHW97] acrescentam que a ARS pode também ser aplicada ao contexto de aprendizagem *online*, mais propriamente, a “redes de aprendizagem”.

2.2.1 Conceito de ARS

A ARS consiste no mapeamento e posterior análise das relações e da troca de informações entre pessoas, grupos ou organizações, não só criando uma representação visual, mas também permitindo uma análise matemática dessas interações.

Os conceitos matemáticos que estão na base da ARS dizem respeito à teoria de grafos, através da qual é possível obter uma representação formal de uma rede, bem como a atribuição de valores às suas propriedades. Através da ARS, torna-se possível interpretar os dados referentes a uma rede social, podendo também obter-se os valores correspondentes aos indicadores que permitem a distinção entre várias redes sociais. A utilização de conceitos ligados à sociologia permite complementar os indicadores matemáticos e a ilustração da rede, atribuindo significados às relações entre os seus participantes.

Um dos aspectos mais importantes das ARS, segundo Laranjeiro (2008), é o facto de permitir a observação das relações entre pessoas que se encontram numa determinada rede social. A diferença principal entre os dados obtidos através das ciências sociais e através da ARS é que, enquanto que os primeiros são obtidos através do estudo individual dos participantes, os segundos obtêm-se através da análise de pares de indivíduos, ou seja, da análise das relações entre indivíduos (Borgatti & Everett, 1997[BE97]; Quiroga, 2005 [QA05]).

2.2.2 Redes de 1-Modo

Nesta secção, a representação das redes de 1-Modo é efectuada através de matrizes. A matriz mais simples, utilizada na representação de redes de 1-Modo, é a matriz binária. Esta matriz representa as relações através de um de dois valores possíveis, 0 ou 1. Caso exista relação entre duas entidades, no espaço correspondente a essa relação é colocado o valor 1. Caso contrário, esse espaço é preenchido com o valor 0. Convém ainda referir que esta matriz pode ser simétrica ou assimétrica, dependendo se as relações entre entidades são recíprocas ou não, respectivamente.

Um Grafo (G) pode ser definido como um conjunto de vértices (V) e um conjunto de ramos (R), tal que:

$$V(G) = \{v_1, \dots, v_n\}$$

$$R(G) = \{r_1, \dots, r_m\}$$

Onde:

- $\{v_1, \dots, v_n\}$: conjunto de vértices do grafo G;
- $\{r_1, \dots, r_m\}$: conjunto de ramos do grafo G.

A matriz adjacente de G, é uma matriz quadrada de dimensão n-por-n, sendo que na linha i e coluna j estará armazenado o número de ramos entre os vértices $\{v_i, v_j\}$. Conclui-se, assim, que se existir um ramo entre dois vértices, estes são considerados adjacentes.

Quando se pretende ler a matriz que armazena as ligações entre actores, constata-se que, nas linhas, são comparados os actores relativamente às escolhas efectuadas e, nas colunas, são comparados os actores que são alvo de escolha por parte dos restantes.

Para ser possível analisar um grupo de actores num determinado fórum de discussão *online*, registam-se as interacções entre os participantes, através da contagem e armazenamento das mensagens enviadas e recebidas por cada um destes. O primeiro passo prende-se com a construção da matriz responsável por armazenar todas as interacções entre os participantes no fórum. De referir ainda que, nas redes de 1-Modo, a primeira mensagem que é considerada diz respeito à primeira resposta efectuada a uma discussão, não sendo considerada a mensagem que corresponde à abertura da discussão.

Tabela 2.1: Disposição de mensagens de um fórum de discussão, em níveis hierárquicos.

0	1	2	3	4
A				
	B			
		C		
		D		
	E			
		B		
			A	
				D
		B		
	E			

Como se pode verificar na tabela 2.1, existe uma mensagem de nível 0, colocada por A. No nível 1 aparecem mensagens de B e E, ou seja, estes actores respondem a A. No nível 2 podem ver-se mensagens dos actores C e D, que respondem a B, e de B, que responde a E. Já no que se refere ao nível 3, apenas consta uma mensagem do actor A, que responde a B. Finalmente, no nível 4 existe igualmente apenas uma mensagem do actor D, que responde a A.

Tabela 2.2: Matriz binária das interacções entre actores.

	A	B	C	D	E
A	0	1	0	0	0
B	1	0	0	0	1
C	0	1	0	0	0
D	1	1	0	0	0
E	1	0	0	0	0

Analisando as colunas da matriz apresentada na tabela 2.2, que diz respeito às mensagens da tabela 2.1, pode-se constatar que os actores C e D não são respondidos por nenhum dos actores que participam neste fórum de discussão. Contudo, analisando as linhas, percebe-se que estes enviam pelo menos uma mensagem a algum dos restantes participantes, não se encontrando, por isso, isolados, ou seja, o número de mensagens em que estão envolvidos, seja enviadas ou recebidas, é superior a 0.

Partindo da análise desta matriz, Laranjeiro (2008) [JBL08] refere que é possível complementar a análise dos dados contidos nessa matriz, através de operações matemáticas simples. Estas operações permitem detectar, por exemplo, os actores mais “populares” ou “influentes” no fórum de discussão em causa, através da soma dos valores que se encontram em cada linha ou em cada coluna que corresponde a um actor. Assim, as colunas cujas somas de valores são mais elevadas dizem respeito aos actores mais “populares”. Por outro lado, as linhas em que as somas tenham valores mais elevados correspondem aos actores com maior nível de “influência”.

Tendo em conta o exemplo apresentado, constata-se, através da análise das linhas, que os actores B e D partilham o mesmo nível de “influência”, tendo, cada um, mensagens de dois actores. Analisando as colunas, percebe-se que os actores A e B partilham o mesmo nível de “popularidade”, recebendo mensagens de 3 actores.

Laranjeiro (2008) [JBL08] refere que, apesar de, geralmente, a diagonal de uma matriz não ter grande relevância na ARS, estes valores podem ser úteis, dependendo das “necessidades de análise”. O mesmo autor acrescenta que, num determinado fórum de discussão, um actor pode responder a si próprio, com o objectivo de “rectificar ou reforçar a sua mensagem”.

Até agora consideramos o caso em que se pretendia analisar apenas a existência de ligações entre actores num determinado fórum de discussão *online*. No entanto, a análise pode ser focada também na quantidade de mensagens trocadas entre os actores, ou seja, citando Laranjeiro (2008), na “intensidade das relações”.

Assim, cada elemento (i,j) da matriz de interacções vai ser preenchido com o número de mensagens enviadas de i para j , caso existam ligações nesse sentido, ou então irá conter o valor 0, se não existirem ligações de i para j .

Tabela 2.3: Matriz binária das interacções entre actores.

	A	B	C	D	E
A	0	1	0	0	0
B	1	0	0	0	2
C	0	1	0	0	0
D	1	1	0	0	0
E	2	0	0	0	0

Comparando a matriz apresentada na tabela 2.3 com aquela que foi mostrada na tabela 2.2, pode observar-se, por exemplo, que a ligação de B para E comporta duas mensagens, algo que não se sabia quando se analisa a matriz binária.

Tal como na matriz anterior, também nesta se pode analisar situações de reciprocidade entre actores. Por exemplo, o actor E envia duas mensagens ao actor A mas este não envia nenhuma mensagem ao actor E. Já no caso das ligações entre A e B, pode-se verificar que A envia uma mensagem a B e este retribui com o envio de uma mensagem.

2.2.3 Redes de 2-Modos

No entanto, Laranjeiro (2008) [JBL08] refere que é importante que ocorra uma análise do ponto de vista das discussões, tendo em conta quem começa uma determinada discussão e quem participa nela. Esta análise pode, segundo o mesmo autor, "ampliar as possibilidades de observação". Conclui-se, assim, que a melhor representação para este tipo de análise obtém-se através do recurso às redes de 2-Modos.

As redes de 2-Modos são representadas por uma matriz incidente n-por-m $M(G)$ em que:

$$m_{ij} > 1, \text{ se } v_i \text{ é um ponto de } r_j$$

Onde:

- m_{ij} : posição (i,j) da matriz m ;
- v_i : vértice i ;
- r_j : ramo j .

A condição apresentada refere que se um vértice i for um ponto do ramo j , conclui-se que i e j são incidentes. Assim, quando dois vértices quaisquer são ligados por um ramo, são denominados como incidentes a esse ramo. Cada um dos campos (i, j) da matriz vai ser preenchido com um valor maior que 0, caso o actor i participe na discussão j , ou com o valor 0, no caso oposto.

Tabela 2.4: Matriz de representação das participações em discussões.

	D1:A	D2:B	D3:C
A	2	1	1
B	1	3	1
C	0	2	2
D	1	0	1
E	0	1	1

Na tabela 2.4, os actores são representados por A, B, C, D e E e as discussões são representadas por D1:A, D2:B e D3:C. No que diz respeito às discussões, a sua representação tem incluída a sua designação e o actor que a iniciou. Por exemplo, a discussão D1:A foi iniciado pelo actor A. Através da observação desta tabela e analisando as linhas, pode-se constatar que, por exemplo, o actor D participa nas discussões D1:A e D3:C, enquanto que o actor E participa nas discussões D2:B e D3:C. Tomando em consideração as colunas, percebe-se que, por exemplo, a discussão D3:C tem participações de todos os actores, enquanto que na discussão D1:A apenas participam 3 dos 5 actores, nomeadamente, A, B e D.

Tal como para o caso das redes de 1-Modo, neste tipo de redes também se pode tirar partido da soma dos valores que compõem as linhas e colunas da matriz para obter dados adicionais de análise. Quando se somam os valores das linhas, obtém-se os actores com maior participação no conjunto de todas as discussões. Caso a opção seja somar as colunas, são obtidas as discussões com maior número de participações. Tendo em conta o exemplo da tabela 2.4, pode-se concluir que o actor B é aquele que tem mais mensagens no conjunto de todas as discussões e a discussão D2:B é a que tem maior número de participações.

2.2.4 Níveis de Análise

Como refere Laranjeiro (2008) [JBL08], na análise de uma rede social, podem utilizar-se três abordagens: individual, grupal e global.

Na primeira abordagem o objectivo é analisar cada indivíduo, tendo em conta as suas ligações a outros membros e o seu posicionamento face aos restantes membros.

A segunda abordagem tem como objectivo a análise de grupos e sub-grupos, ou seja, das áreas da rede com maior conectividade entre vários dos seus membros. Para que um conjunto de membros de uma rede possa ser considerado um grupo, tem que ter no mínimo três membros e possuir ligações, em ambos os sentidos, entre todos eles.

A terceira e última abordagem pretende analisar a rede como um todo, tendo em conta todas as ligações presentes na rede e todos os membros responsáveis por estas, destacando os membros com um papel central na rede e aqueles que têm um nível de importância muito baixo, ou seja, aqueles que são periféricos na rede.

2.3 Tecnologias adoptadas e justificação

2.3.1 Linguagem de desenvolvimento da aplicação

A linguagem de programação utilizada para desenvolver a aplicação é a linguagem Java. Esta possui uma interface que contém grafos correspondentes às interacções entre actores em fóruns de discussão *online*, bem como um conjunto de opções para a visualização desses grafos e ainda um conjunto de informações sobre um determinado actor que se pretenda visualizar, através do clique no nó correspondente a esse actor.

Para isso, é utilizada uma API Java para interfaces gráficas, a Swing. Esta API desenha todos os componentes da interface, ao contrário da maioria das outras APIs de interfaces gráficas que delegam essa tarefa no sistema operativo.

Esta API tem a desvantagem de consumir muita memória RAM durante a sua execução. No entanto, tem a enorme vantagem de ter uma aparência muito semelhante, qualquer que seja o sistema operativo em que é executada, tornando-a assim portátil.

2.3.2 Sistema de Gestão de Bases de Dados

O SGBD é responsável por guardar todos os dados referentes aos fóruns de discussão, nomeadamente, os dados dos actores e das suas interacções em cada uma das *threads* dos fóruns de discussão. Estes dados irão, posteriormente, ser utilizados para analisar as interacções entre os diversos actores.

Uma vez que os dados dos fóruns de discussão e respectivos participantes vão ser importados da base de dados do Moodle, cujo SGBD é o MySQL, é este o SGBD utilizado na execução do *script* de recolha de dados (explicado na secção 4.2), responsável por obter os dados das interacções entre actores num determinado fórum de discussão no Moodle. O MySQL é também utilizado para importar os dados dos actores para a interface da aplicação que vai ser desenvolvida.

2.3.3 Linguagem de desenvolvimento web

A linguagem de programação em ambiente *web* é utilizada para o desenvolvimento do módulo responsável por incorporar a aplicação no Moodle.

A escolha poderia recair entre uma das três linguagens seguintes (baseadas no conceito de *scripts* do lado do servidor):

- PHP;
- ASP;
- JSP.

2.3.3.1 PHP

PHP é uma linguagem de *script* que gera o código HTML das páginas que são apresentadas ao utilizador. Esta linguagem cria páginas dinâmicas que interagem com uma ou mais bases de dados alojadas num determinado servidor.

2.3.3.1.1 Vantagens da linguagem PHP

A linguagem PHP tem as seguintes vantagens:

- Completamente gratuita: *software* de instalação grátis e facilmente descarregável da Internet;
- Multi-plataforma: funciona em qualquer plataforma, desde que nela seja possível instalar um servidor *web*. Exemplos de alguns sistemas operativos compatíveis com PHP: Windows, Linux, Unix, Solaris, etc;
- Fácil de aprender: baseia-se em três linguagens: Perl, Java e C. A maior parte dos programadores *web* conhecem pelo menos uma destas três linguagens, o que facilita a sua aprendizagem;
- Compatibilidade com muitos SGBDs: é compatível com a maior parte dos SGBDs, como por exemplo MySQL, PostgreSQL, Oracle e SQL Server, entre muitos outros compatíveis com o padrão ODBC;
- Velocidade no acesso aos dados: rápido acesso aos dados que se encontram em bases de dados num servidor. Raramente deixa o servidor lento quanto está a aceder aos dados;
- Segurança: o código fonte é preservado, uma vez que o Servidor retorna somente código HTML;
- Código *open source*: muitos módulos de código estão disponíveis na *web* e podem ser consultados e reutilizados. Isto ajuda claramente quando os programadores têm

dúvidas e, assim, podem consultar código já desenvolvido. Uma razão para este facto é a comunidade PHP, que é enorme e se encontra constantemente em expansão.

2.3.3.1.2 Desvantagens da linguagem PHP

Tem algumas desvantagens, nomeadamente, a parte orientada a objectos ser mais débil, comparando com JSP, mas é pouco relevante esse facto, uma vez que a versão PHP5 melhorou bastante nesse aspecto.

2.3.3.2 ASP

ASP é uma estrutura de programação ou uma *framework* em *script*, embora se diga normalmente que é uma linguagem de programação *web*. Utiliza VBScript, JScript, PerlScript ou Python. Tal como PHP, processa o código do lado do servidor para geração de conteúdo dinâmico na *web*.

2.3.3.2.1 Vantagens da linguagem ASP

A linguagem ASP tem as seguintes vantagens:

- Gratuita para Windows;
- Aprendizagem: baseada em Visual Basic. Caso se saiba programar nesta linguagem, torna-se fácil transitar para ASP;
- Segurança: partilha da mesma vantagem que o PHP no que respeita à segurança.

2.3.3.2.2 Desvantagens da linguagem ASP

No que diz respeito a desvantagens no desenvolvimento de aplicações em ASP, destacam-se as seguintes:

- Mono-plataforma: só pode ser executado em Windows;
- Licenças caras: utilizar ASP para manipular bases de dados implica ter licença do Microsoft Office, para utilizar Access, ou licença para Microsoft SQL Server, para o utilizar;
- Hospedagem mais cara: hospedar código ASP é, na maioria dos servidores, mais caro do que hospedar código PHP;
- Poucos SGBDs compatíveis: apenas é compatível com SGBDs Windows (Access, Microsoft SQL Server);

- Dependência de futuras modificações: num futuro, próximo ou mais distante, ASP pode ser substituído por outra linguagem, forçando os programadores a “esquecer” tudo o que aprenderam com ASP e aprender uma nova linguagem. Isto já aconteceu, em parte, com a actualização de ASP para ASP.NET, forçando mudanças substanciais na programação para a *web*;
- Código fechado: ao contrário de PHP, não há código ASP disponível na *web* para consulta ou reutilização.

2.3.3.3 JSP

JSP é uma tecnologia padrão, baseada em templates para *servlets*. Estas *servlets* são classes na linguagem de programação Java que, dinamicamente, processam requisições e respostas, proporcionando, desta forma, novos recursos aos servidores. O mecanismo que a traduz é embutido no servidor.

2.3.3.3.1 Vantagens da linguagem JSP

JSP tem as seguintes vantagens:

- Completamente gratuita, tal como PHP;
- Multi-plataforma: partilha da mesma vantagem que PHP no que concerne à execução em diversas plataformas;
- Fácil de aprender: tal como PHP, também se baseia na linguagem Java, sendo esta uma das linguagens mais conhecidas pelos programadores e com mais documentação na *web*;
- Compatibilidade com muitos SGBDs: é compatível com a maior parte dos SGBDs, tal como PHP;
- Segurança: partilha da mesma vantagem que ambas as linguagens anteriores no que respeita à segurança;
- Código *open source*: tal com PHP, muitos módulos de código JSP estão disponíveis na *web* e podem ser consultados e reutilizados. Tal como a comunidade PHP, a comunidade Java na *web* também é enorme.

2.3.3.3.2 Desvantagens da linguagem JSP

Quanto a desvantagens na utilização de JSP, destaca-se a baixa velocidade no acesso aos dados, tendo como comparação PHP.

2.3.3.4 *Comparações*

Analisando vantagens e desvantagens das três linguagens analisadas anteriormente, pode concluir-se que PHP é a linguagem mais adequada para o projecto em causa. Comparando com ASP, tem várias vantagens, nomeadamente:

- Custo;
- Compatibilidade com muitos sistemas operativos;
- Compatibilidade com muitos SGBDs;
- Baixa probabilidade de haver modificações acentuadas na sintaxe nos próximos anos;
- Código *open source*.

No que toca a desvantagens de PHP relativamente a ASP, não conhecemos qualquer desvantagem significativa.

Comparando com JSP, as vantagens da utilização de PHP são praticamente as mesmas, apenas sendo melhor na rapidez de acesso a bases de dados de um determinado servidor. Tem a desvantagem de ter uma parte orientada a objectos mais débil, mas a vantagem em relação a JSP, mencionada acima, sobrepõe-se a esta desvantagem.

Não obstante todas estas considerações anteriores e a conclusão de que a linguagem PHP é a mais indicada para o objectivo pretendido, é um dado adquirido que o Moodle, onde vai ser incorporada a aplicação, foi desenvolvido em PHP. Logo, esta teria que ser, em qualquer das situações a linguagem utilizada para a implementação do módulo pretendido.

De qualquer maneira, com todas as considerações sobre as 3 linguagens de desenvolvimento *web* que foram alvo de análise, pode concluir-se que o Moodle foi implementado na melhor linguagem que permite o desenvolvimento de páginas dinâmicas na *web*.

2.4 Tecnologias existentes para ARS

Foram estudadas as quatro ferramentas existentes, mais relevantes no âmbito da ARS: NetDraw, Graphviz, UCINet e Pajek.

2.4.1 UCINet

A ferramenta UCINet tem como objectivo principal efectuar a ARS em redes de 1-Modo e de 2-Modos, tendo em conta somente o cálculo das métricas. Esta ferramenta possui uma ampla abrangência de métricas, procedimentos e operações, auxiliada pela grande variedade de opções de formatos de importação e exportação de ficheiros, tornando-a na melhor ferramenta quando apenas são tomadas em conta estes factores para a ARS. No entanto, esta ferramenta não

possibilita a visualização do grafo relativo às interacções nem pode ser embebida num LMS. Para colmatar a falha relativa à visualização do grafo, pode ser utilizada em conjunto com as ferramentas NetDraw e Pajek, entre outras.

2.4.2 NetDraw

Em relação à ferramenta NetDraw, esta é utilizada para efectuar o *layout* de um grafo correspondente às interacções entre participantes. No entanto, não possibilita, nas suas funcionalidades, o cálculo de métricas para complementar a análise do grafo nem a possibilidade de embeber essa visualização num LMS.

2.4.3 Graphviz

Em relação à ferramenta Graphviz, esta permite a visualização de um grafo relativo às interacções entre participantes, após a leitura do ficheiro correspondente. Fornece também a possibilidade de colocar os grafos em ambiente *web*, o que, no caso deste trabalho, seria útil para colocar a representação das interacções num LMS. Contudo, tem um ponto fraco referente aos indicadores numéricos de análise, uma vez que estes não considerados nesta ferramenta.

2.4.4 Pajek

A ferramenta Pajek tem o objectivo de permitir a completa análise de uma rede, segundo os conceitos ligados à ARS. Tem como grande aliciente a possibilidade de conjugar os dois aspectos importantes quando se fala em ARS: a análise matemática e a análise visual. De facto, possui algoritmos de *layout* poderosos e complementa esta visualização através do cálculo de um conjunto de métricas, embora em menor número, comparando com a ferramenta UCInet. Tem como desvantagem, no contexto deste trabalho, não ter disponível a funcionalidade para embeber a visualização do grafo e as métricas de análise num LMS.

2.4.5 Comparações

Analizando as quatro ferramentas, referidas anteriormente, verificamos que nenhuma delas cobre os três objectivos principais deste trabalho, nomeadamente, *layout* do grafo das interacções, cálculo dos indicadores numéricos de análise e possibilidade de embeber o grafo e as métricas num LMS. De facto, algumas delas cobrem dois objectivos, como é o caso da ferramenta Pajek, que faz o *layout* do grafo e calcula as métricas, embora não permita embeber estes resultados num LMS.

Posto isto, concluímos que a opção possível seria desenvolver uma aplicação que conjugasse todos os objectivos pretendidos.

3 Representação de Grafos e Indicadores Numéricos de Análise

Os conceitos, descritos a seguir, dizem respeito aos indicadores numéricos de análise, ou seja, às métricas utilizadas para uma melhor compreensão das interações entre os actores que se encontram num determinado fórum de discussão *online*. Estas métricas são o Grau de Centralidade, a Densidade, o Índice de Centralização e os Cliques, tanto na sua vertente mais estrita como numa vertente mais lata.

O último tópico, descrito neste capítulo, prende-se com os conceitos que serviram de base à implementação do algoritmo para o desenho do grafo correspondente a um fórum de discussão *online*.

3.1 Indicadores numéricos de análise

Um dos pontos focados neste capítulo refere-se aos indicadores numéricos de análise ou métricas. Estas métricas revestem-se de uma importância enorme, uma vez que permitem complementar a informação relativa ao fórum de discussão em análise, obtida através da visualização do grafo correspondente.

A visualização do grafo é um aspecto central na análise do fórum em causa, permitindo verificar actores centrais ou actores isolados, actores que recebem mais mensagens ou que, por outro lado, enviam mais mensagens, actores que recebem mais mensagens do que enviam ou o contrário. Todos estes factos são passíveis de ser visualizados aquando da representação do grafo na aplicação. Mas é neste ponto que entram as métricas e a sua importância na complementaridade da análise ao fórum de discussão em causa. Estas métricas permitem avaliar quantitativamente a importância de cada actor no fórum em causa.

Por exemplo, quando se constata, através da visualização do grafo, que um actor é central, ou seja, que é importante na distribuição de conhecimento pelos restantes actores do fórum, essa constatação é apenas qualitativa. Quer isto dizer que essa constatação é apenas de carácter

objectivo, ou seja, sabe-se que o actor é central mas não se sabem as diferenças de centralidade para os outros actores. É neste ponto que se revela a importância das métricas para complementar essa análise. Através da utilização do Grau de Centralidade, podemos constatar as diferenças entre a “centralidade” do actor mais central e as “centralidades” dos restantes, podendo, assim, concluir quão central é esse actores em relação aos restantes, ou seja, qual vai ser o nível de importância deste, no fórum de discussão em causa.

Nesta secção vão ser descritos os indicadores numéricos de análise utilizados na aplicação, tanto a nível local (Grau de Centralidade), como a nível grupal (Cliques), e ainda a nível global (Densidade e Índice de Centralização).

3.1.1 Análise Local

3.1.1.1 Grau de Centralidade

Este indicador está ligado à análise local dos fóruns de discussão, dizendo respeito, por isso, a cada um dos actores, individualmente. Este indicador é calculado, tendo por base a seguinte fórmula:

$$C_i^{GC} = \sum a_{ij}$$

Onde:

- C_i^{GC} : Grau de Centralidade do actor i ;
- a_{ij} : vértice j adjacente a i .

Esta fórmula calcula o Grau de Centralidade de um actor, através do somatório dos actores adjacentes ao actor em questão.

O grau de centralidade mínimo que um actor pode ter é 0 e, neste caso, é considerado um actor isolado, ou seja, não interage com nenhum dos restantes actores presentes no fórum em análise.

Como refere Laranjeiro (2008) [JBL08], este indicador revela quais são os actores que têm maior relevância para os restantes, percebendo-se, assim, quais são os actores com maior relevância no desenrolar da actividade do fórum de discussão. Laranjeiro (2008) [JBL08], refere ainda que estes actores são aqueles que transmitem mais conhecimento aos restantes, e são vistos por estes como veículos de informação de qualidade, que lhes vão permitir desenvolver o seu conhecimento na área temática relacionada com o fórum em questão.

Este Grau de Centralidade, analisado anteriormente, diz respeito a uma análise mais geral, ou seja, analisa o número de actores ligados a um actor específico, não analisando se esses actores enviam mensagens a esse actor ou se é este que lhe envia mensagens a eles.

Representação de Grafos e Indicadores Numéricos de Análise

Para uma análise mais concreta nesse aspecto, existem dois indicadores, o Grau de Centralidade de Entrada e de Saída, que calculam, respectivamente, a quantidade de ligações dirigidas para o actor em causa e a quantidade de ligações que saem desse actor para outros.

Os actores com Grau de Centralidade de Entrada superior são aqueles que têm mais ligações de entrada e, por isso, como refere Laranjeiro (2008) [JBL08], são considerados mais “populares” ou com mais “prestígio”. Já aqueles que tenham um Grau de Centralidade de Saída maior são aqueles que têm mais ligações de saída, sendo, por isso, mais “influentes”, trocando mais informação com os restantes actores.

De referir ainda que estes dois indicadores apenas são considerados na rede de 1-Modo, já que apenas esta consiste num grafo direccionado.

Até agora apenas foi explicado o conceito de Grau de Centralidade tendo por base o número de interacções que um actor possui, sem se considerar a quantidade de mensagens que pode englobar cada uma dessas interacções consideradas.

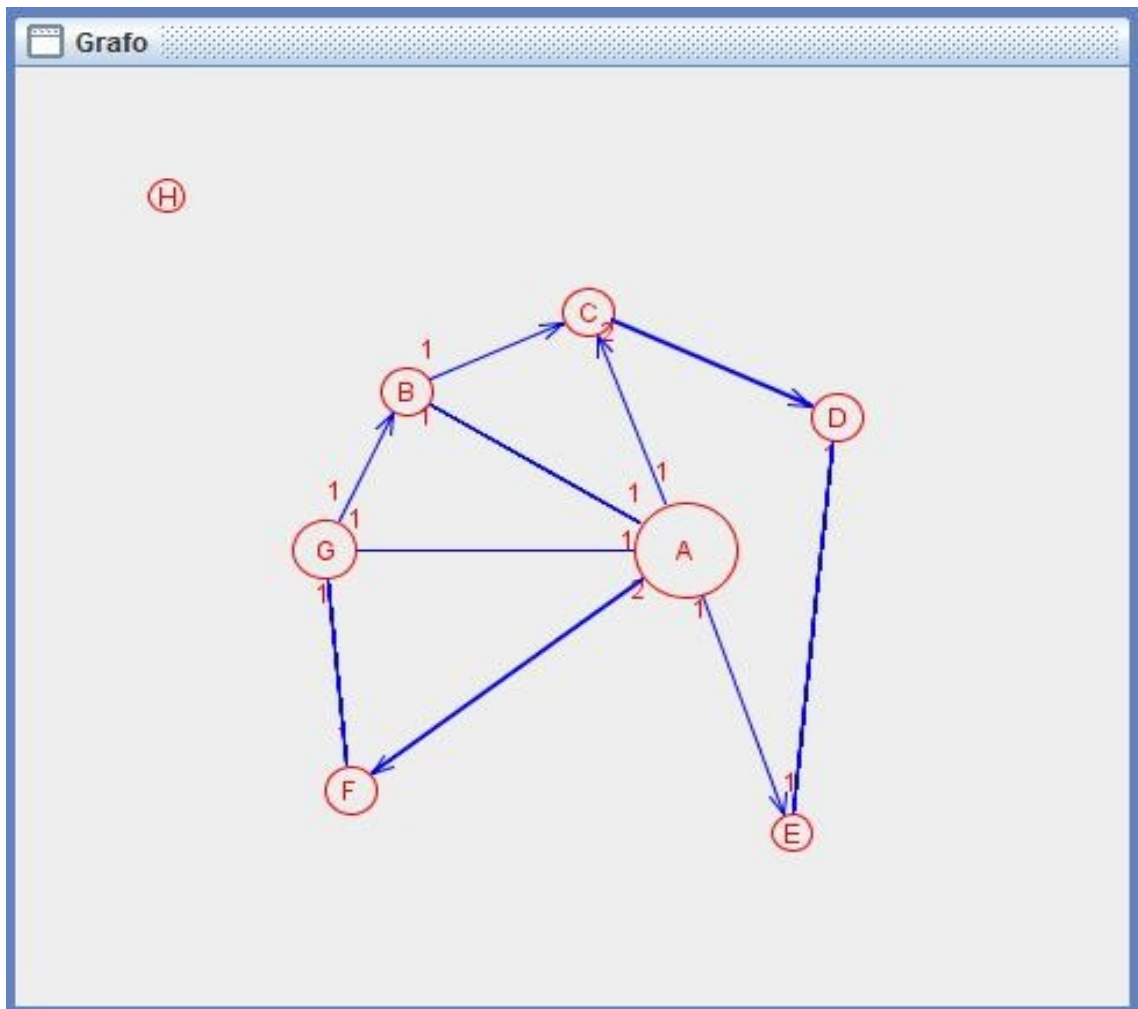


Figura 3:1: Exemplo de um grafo representado na interface da aplicação

Mas o Grau de Centralidade pode também ser analisado quantitativamente, ou seja, analisando cada interacção como sendo uma ou mais mensagens que o actor recebe ou envia. Esta análise utiliza os pesos das ligações para um determinado actor, ao contrário da análise qualitativa, explicada anteriormente, que soma apenas as ligações a esse actor. Os Graus de Centralidade de Entrada e de Saída também podem ser considerados tendo em conta os pesos das ligações a um determinado actor, sendo o seu cálculo em tudo idêntico ao que já foi explicado anteriormente sobre estes dois indicadores, variando apenas no aspecto de, agora, se considerar os pesos das ligações entre os actores, em vez de se considerar simplesmente o facto de haver ligações entre eles.

Tendo por base o grafo apresentado na figura 3.1, obtém-se a seguinte tabela de Graus de Centralidade.

Tabela 3.1: Grau de Centralidade, tendo em conta a existência de ligações.

Actores	GC	GCE	GCS
A	5	2	5
B	3	2	2
C	3	2	1
D	2	2	1
E	2	2	1
F	2	2	1
G	3	2	3
H	0	0	0

Na tabela acima apresentada (Tabela 3.1), pode-se constatar que, tendo em conta apenas o número de ligações a outros actores, o A é aquele que tem mais actores ligados a si (5), no que diz respeito a ligações de saída de informação. Quanto à entrada de informação, todos os actores têm duas ligações de entrada com outros actores, à excepção de H. Este actor não tem nenhuma ligação com os restantes actores e, por isso, percebe-se que é um actor isolado. Isto quer dizer que o actor H não tem contacto com qualquer um dos outros actores no fórum de discussão em análise.

Tabela 3.2: Grau de Centralidade, considerando o peso das ligações.

Actores	GC	GCE	GCS
A	8	2	6
B	4	2	2
C	4	2	2
D	4	3	1
E	3	2	1
F	3	3	1
G	5	2	3
H	0	0	0

Na Tabela 3.2, são apresentadas as três variantes do Grau de Centralidade, tendo em conta o número efectivo de mensagens trocadas entre os actores. No que diz respeito ao número de mensagens recebidas, destacam-se os actores D e F com 3 mensagens. Quanto às mensagens enviadas, é o actor A quem se destaca, com 6 mensagens, seguido de longe pelo actor G com 3 mensagens enviadas. O actor H não recebe nem envia qualquer mensagem aos outros actores, sendo, por isso, um actor isolado.

Por um lado, conclui-se que o actor A é o mais “influente”, ou seja, é aquele que transmite mais informação para os restantes actores do fórum, podendo ser visto por estes como um bom veículo de conhecimento para os ajudar na sua aprendizagem. Por outro lado, constata-se também que os actores D e F são os mais “populares”, uma vez que são aqueles que recebem mais mensagens dos restantes, embora com uma diferença reduzida em comparação com os outros actores.

3.1.2 Análise Grupal

3.1.2.1 Cliques

Os cliques são indicadores que permitem localizar grupos de actores, dentro de um determinado fórum de discussão, tendo, por isso, uma índole grupal. Um clique é um conjunto de três ou mais nós de um grafo que têm todas as ligações possíveis entre eles.

Neste projecto, são considerados dois tipos de Cliques, nomeadamente, Cliques em “sentido estrito” e Cliques em “sentido lato”.

Um Clique em sentido estrito é detectado se cada um dos actores dentro desse grupo enviarem e receberem mensagens de/para todos os restantes actores do grupo. Isto permite detectar grupos com maior conectividade entre os actores e analisar esse grupo sob uma perspectiva global.

Em relação a um Clique em sentido lato, este é encontrado se existir uma ligação, quer seja de entrada ou de saída, de um actor para todos os restantes actores do grupo.

3.1.3 Análise Global

3.1.3.1 Densidade

A Densidade de uma rede é um indicador de índole global, uma vez que serve para analisar toda a rede de um determinado fórum, sendo um dos indicadores mais utilizados para análise de redes sociais (ARS).

A fórmula para calcular a Densidade é a seguinte:

$$D = (\sum a_{ij}) \div |V| \times (|V| - 1)$$

Onde:

- D : Densidade da rede;
- $\sum a_{ij}$: soma de todas as ligações existentes na rede;
- V : número de actores na rede.

A Dnsidade é calculada, através da divisão entre a soma de todos as ligações existentes pela número de ligações possíveis na rede (AP) que é igual a:

$$AP = |V| \times (|V| - 1)$$

Uma vez que a densidade vai ser um valor em percentagem, no final do cálculo da fórmula acima descrita, o resultado é multiplicado por 100.

Através do valor obtido com este cálculo da densidade, pode concluir-se se a intensidade de circulação de informação na rede é elevada ou reduzida. Se o valor da Densidade for de 100%, isto quer dizer que a rede é completa, uma vez que existem todas as ligações possíveis dentro do grafo, ou seja, todos os actores interagem com os restantes em ambos os sentidos. Se o valor da Densidade estiver próximo de 100%, conclui-se, tal como refere Laranjeiro (2008) [JBL08], que a rede tem uma intensidade de troca de informação elevada. Se este valor estiver mais afastado dos 100%, Laranjeiro (2008) [JBL08] conclui exactamente o contrário, ou seja, que a rede tem poucas trocas de informação.

A Densidade pode ainda ser calculada tendo por base duas condições diferentes, nomeadamente, considerando a rede completa, com todos os actores de um fórum geral, ou considerando apenas os participantes activos, ou seja, os actores que participam num fórum específico, dentro do geral, sendo aquele que no momento esta a ser alvo de análise. Assim, é possível analisar se a Densidade da rede é alta ou baixa, tendo em conta apenas os actores que, efectivamente, participam com o seu contributo na troca de informação no fórum em análise.

Todas estas considerações dizem respeito ao cálculo da Densidade para a rede de 1-Modo. No que concerne à rede de 2-Modos, uma vez que não existem ligações entre os actores no mesmo conjunto, o número máximo de ligações possíveis obtém-se quando todos os actores num conjunto estão ligados a todos os actores do outro.

Tendo em conta o exemplo apresentado na figura 3.1, a Densidade vai ser apresentada na seguinte tabela.

Tabela 3.3: Densidade do grafo para participantes activos ou rede completa.

ParticipantesActivos	RedeCompleta
24%	19%

Como se pode constatar, e é normal que assim seja, o valor da Densidade para os participantes activos é superior à densidade para a rede completa. Isto deve-se ao facto do actor J se encontrar isolado e, por isso, não entra no cálculo da Densidade para os participantes activos, fazendo com que esta seja um pouco mais alta que a Densidade calculada para a rede completa.

Os valores calculados são baixos, em relação ao ideal que seria 100%, o é explicado pela baixa actividade entre os participantes activos no fórum, havendo apenas 24% das interacções possíveis entre eles. Conclui-se, portanto, que este fórum tem uma baixa actividade entre os actores inscritos nele.

3.1.3.2 Índice de Centralização

O Índice de Centralização é um indicador numérico de análise global. Como refere Laranjeiro (2008) [JBL08], este indicador analisa a centralidade de um actor, em relação aos restantes. Tal como a Densidade, é expresso em percentagem.

Este indicador calcula-se através da seguinte fórmula:

$$IC = \sum [c_* - c_i] \div \max \sum [c_* - c_i]$$

Onde:

- IC : Índice de Centralização da rede;
- c_* : Valor maior do Grau de Centralidade de entre todos os actores da rede;
- c_i : Valor do Grau de Centralidade dos restantes actores.

O Índice de Centralização é calculado através de vários passos. O primeiro passo é somar a diferença entre o maior Grau de Centralidade e o Grau de Centralidade dos restantes actores. Depois disso, é somada a diferença máxima entre o maior Grau de Centralidade e os restantes Graus de Centralidade, admitindo que se trata do máximo possível, ou seja, o maior Grau de Centralidade é igual ao número de actores menos uma unidade e os restantes Graus de Centralidade têm o valor 1. Finalmente, dividem-se as duas somas efectuadas nos primeiros passos e o resultado final é multiplicado por 100, para dar um valor em percentagem.

O Índice de Centralização é de 100% se, tal como refere Laranjeiro (2008) [JBL08], a rede em análise possuir um actor central com ligações, em ambos os sentidos, para todos os actores e

os restantes terem apenas uma ligação, em ambos os sentidos, para o actor central. Este tipo de rede chama-se Rede em Estrela e é exemplificada na figura 3.2.



Figura 3:2: Rede em Estrela (Índice de Centralização igual a 100%).

Por outro lado, se o Índice de Centralização é de 0%, Laranjeiro (2008) [JBL08] conclui que não existe um actor central, ou seja, todos os actores possuem o mesmo número de ligações para os restantes. A figura 3.3 demonstra um exemplo deste caso.



Figura 3:3: Rede sem um actor central (Índice de Centralização igual a 0%).

Tal como na Densidade, o Índice de Centralização pode ser calculado não só para a rede completa, mas também apenas para os participantes activos, possibilitando, esta última opção, uma análise mais realista acerca da centralização de um determinado fórum sendo apenas considerados os actores que realmente interagem dentro deste.

Existem duas variantes do Índice de Centralização, que são o Índice de Centralização de Entrada e o Índice de Centralização de Saída.

O Índice de Centralização de Entrada calcula a centralização da rede tendo apenas em conta as mensagens recebidas. Para o seu cálculo, utiliza o Grau de Centralidade de Entrada.

Representação de Grafos e Indicadores Numéricos de Análise

Ao contrário do anterior, o Índice de Centralização de Saída tem em conta as mensagens enviadas para calcular a centralização da rede, utilizando para o seu cálculo o Grau de Centralidade de Saída.

Tendo em conta, tal como para a Densidade, o fórum representado na figura 3.1, a tabela seguinte mostra os valores do Índice de Centralização desse fórum.

Tabela 3.4: Índice de Centralização, para a rede completa e participantes activos.

	IC	ICE	ICS
RedeCompleta	50%	60%	48%
ParticipantesActivos	57%	74%	52%

Tendo por base os valores obtidos na tabela 3.4, constata-se que o Índice de Centralização, do fórum representado na figura 3.1, é relativamente alto, tal como já se podia constatar ao analisar os valores dos Graus de Centralidade, com o actor A em destaque, à semelhança do que já se podia constatar na visualização do grafo.

Outra detalhe que vale a pena realçar é o facto de o Índice de Centralização ser superior na sua vertente de Entrada, tal como também facilmente se constata na tabela respeitante aos Graus de Centralidade no que respeita as mensagens trocadas, vendo-se que são muito mais as respostas enviadas do que aquelas que são recebidas pelos actores.

3.2 Redes de 1-Modo

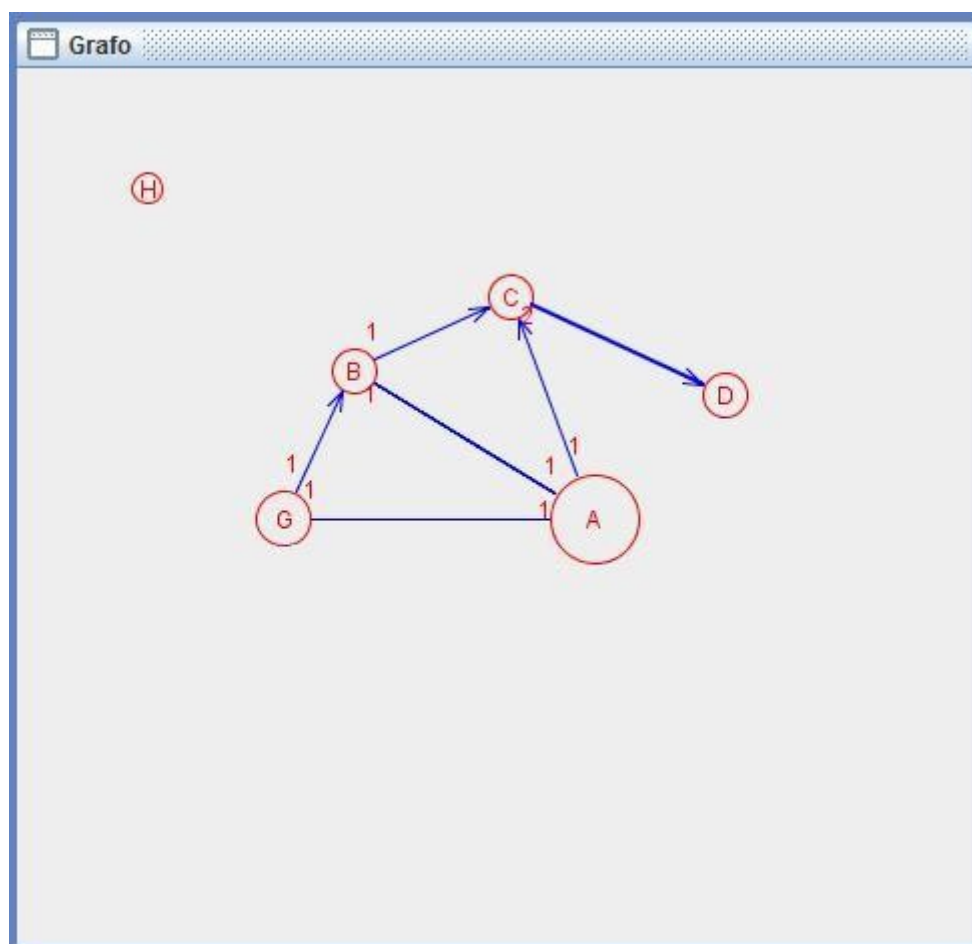


Figura 3:4: Grafo sob a forma de rede de 1-Modo

Como refere Laranjeiro (2008) [JBL08], as redes de 1-Modo representam as interações entre actores num determinado fórum de discussão *online*.

Laranjeiro (2008) [JBL08] define alguns pressupostos na representação dos nós e respectivas ligações, por forma a ser mais fácil a sua compreensão e interpretação.

Em relação os nós, estipulou o seguinte:

- Os nós estão legendados com a designação atribuída ao respectivo actor;
- Os nós são diferenciados por tamanho, sendo este definido pelo número de mensagens.

Em relação às ligações, definiu o seguinte:

- A espessura de cada um das linhas é definida pelo número de mensagens trocadas entre os dois actores unidos por uma determinada linha, sendo mais espessas quanto maior for o número de mensagens trocadas entre ambos;
- Legendagem das ligações com o respectivo peso, ou seja, com o número de mensagens enviadas de um actor para o outro. O número de mensagens enviadas é aquele que está mais próximo do actor que as enviou;
- Quando um actor responde a si próprio, esta ligação é representada com um arco para ele próprio, também designado por laço;
- Se dois actores trocam mensagens entre si, a ligação não irá ter direcção. No caso de apenas um deles enviar mensagens ao outro, essa ligação irá ter a direcção do actor para onde foram encaminhadas as mensagens.

Analisando a Figura 3.4, podemos tirar diversas conclusões. No que diz respeito ao tamanho dos nós, distingue-se o actor A com maior número de mensagens trocadas. Através da legendagem dos nós, associa-se um actor a um determinado nó e a todos os indicadores inerentes a esse nó. Já em relação às ligações, é, igualmente, o actor A aquele que tem mais ligações e, por isso, é o actor central da rede. Em contraponto, o actor H não estabelece nenhuma ligação com outro actor, sendo, por isso, um actor isolado. Em relação ao direccionamento das setas, consegue-se constatar que, por exemplo, entre A e B a troca de mensagens faz-se em ambos os sentidos. Já entre B e C, a ligação apenas é estabelecida de B para C, ou seja, C não envia mensagens a B, acontecendo apenas o contrário.

3.3 Redes de 2-Modos

Para este tipo de rede, tal como para as redes de 1-Modo, Laranjeiro (2008) [JBL08] definiu propriedades para a sua representação, tendo em conta os nós e os ramos.

Em relação aos nós, definiu as seguintes propriedades:

- São constituídos por dois conjuntos: o conjunto (Q) das discussões (representada por quadrados) e o conjunto (C) dos actores que participam nessas discussões (representados por círculos). Os vértices deste novo tipo de grafo são dados por:

$$V = Q \cup C$$

- As legendas dos círculos assinalam o actor correspondente. Nos quadrados, a legenda corresponde à designação da discussão e ao actor que a iniciou, segundo a sintaxe "Discussão:Actor".

Em relação aos ramos, definiu a seguinte propriedade:

- A espessura das linhas é definida pelo número de participações que um actor tem numa discussão, ou seja, as participações que uma discussão recebe de um determinado actor. Estas linhas são mais espessas quanto maior for o número de participações de um actor numa determinada discussão.

Este tipo de redes não foi implementado na aplicação por falta de tempo.

3.4 Conceitos importantes do algoritmo de desenho do grafo

Tal como refere Figueira (2009) [FAR09], existe uma vasta literatura sobre o desenho automático de grafos, tendo sido desenvolvidos, nos últimos 30 anos, muitos algoritmos de *layout* baseados na teoria de grafos. Figueira (2009) [FAR09] define alguns princípios que deverão servir de base ao algoritmo de desenho automático de grafos, no sentido de proporcionar um entendimento humano mais confortável e rápido do grafo:

- Distribuição dos nós, evitando sobreposições;
- Actores centrais devem ter tendência para ser colocados mais perto do centro;
- Minimização do cruzamento de nós e arestas;
- Agrupamento de cliques;
- Tendência para uma rede com elevada densidade se espalhar equitativamente pelo espaço reservado ao grafo.

No entanto, o cumprimento de grande parte destes requisitos é extremamente complicado quando uma determinada rede em análise tem um número elevado de actores e uma elevada densidade. De facto, percebe-se que a distribuição dos nós, para que não haja sobreposições, é complicada, porque, para um número elevado de actores, existe dificuldade em colocá-los num espaço limitado, garantindo este princípio. Da mesma forma, evitar ou minimizar a ocorrência de cruzamentos entre nós e arestas, com elevado número de actores e uma densidade elevada, reveste-se de uma complexidade enorme, podendo, em vários ocasiões, não existir uma solução óptima para este princípio. A pesquisa de cliques, para a mesma situação, é outro princípio com elevada complexidade de computação, uma vez que vão ser analisados todas as interacções de todos os actores, exigindo muita memória/tempo para calcular este indicador numérico. Tendo em conta a mesma situação já referida, espalhar equitativamente a rede pelo espaço destinado ao grafo, minimizando as distâncias entre nós e cliques e distanciando os nós de cada cliques com o mesmo espaço, é igualmente caracterizado por envolver uma extensa computação, já que terão de ser permutados os nós de cada clique e verificar, a cada permutação, qual a distância aos restantes nós fora do clique, obtendo-se a distância mínima dos restantes nós ao clique considerado.

4 O Protótipo

Neste capítulo são apresentados os aspectos ligados à implementação da aplicação pretendida para a análise de interacções entre actores no contexto de fóruns de discussão *online*. Inicialmente, é explicada a arquitectura da aplicação, sendo apresentado um esquema representativo. Posteriormente, é explicado o *script* de recolha de dados, através da apresentação do seu pseudo-código, bem como de um exemplo da sua execução. Segue-se a explicação do “baixo nível” da implementação, contendo os seus pontos mais relevantes. O último destes, diz respeito ao algoritmo de representação do grafo das interacções entre actores, sendo, neste ponto, explicado o algoritmo e apresentado um exemplo da sua execução. Ainda neste capítulo, é efectuada uma visita guiada à interface da aplicação, sendo explicados todos os pontos que a constituem. Finalmente, é explicada a *applet* implementada, com vista a possibilitar a utilização da aplicação em ambiente *web*.

4.1 Arquitectura da aplicação

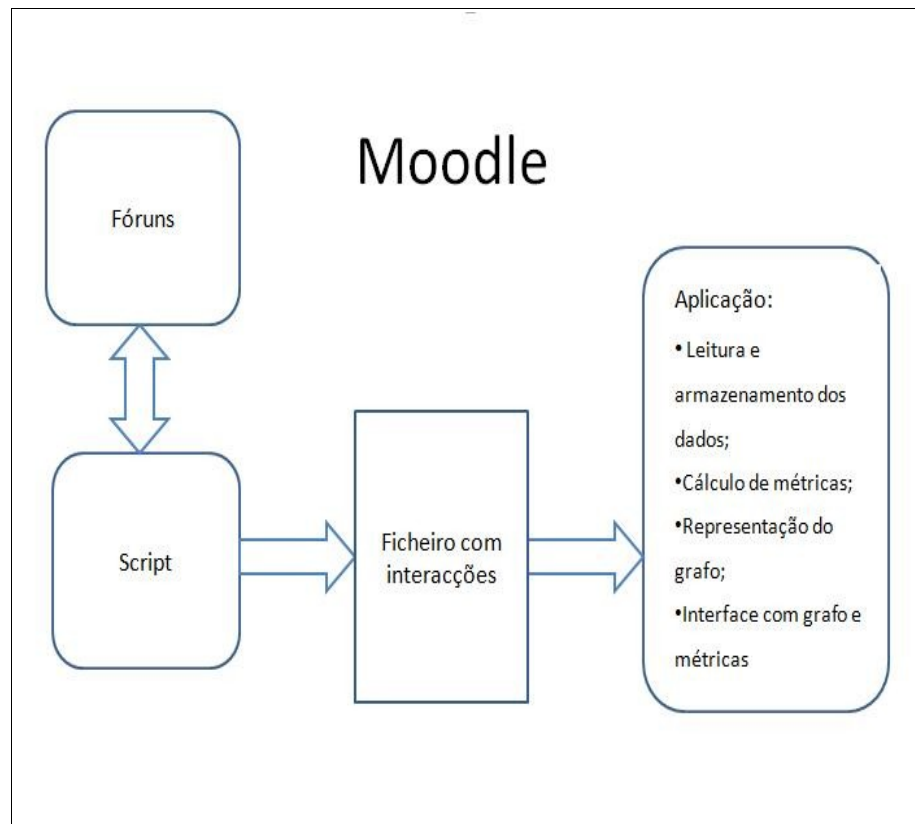


Figura 4:1: Arquitectura da aplicação.

Como mostra a figura 4.1, a arquitectura da aplicação desenvolvida tem três importantes elementos. Um deles é o Moodle, que corresponde ao LMS onde se pretende embeber a interface da aplicação. Outro dos componentes da arquitectura da aplicação é o *script* de recolha de dados, que recebe, do Moodle, as interações dos alunos nos fóruns de uma determinada disciplina e guarda-as num ficheiro. O último componente da arquitectura apresentada é a interface da aplicação, que lê e interpreta os dados do ficheiro referido e utiliza esses dados para o cálculo de indicadores numéricos e para o desenho do grafo correspondente, facultando, assim, informação suficiente para se proceder à análise de um determinado fórum, tendo em conta os conceitos de ARS.

4.2 Script de recolha de dados

O *script* de recolha de dados baseia-se na obtenção das mensagens relativas às interações de actores nos diferentes tópicos de cada um dos fóruns de discussão das várias disciplinas leccionadas que se encontram no Moodle. Este *script* serve de base à posterior análise dessas interações, uma vez que gera um ficheiro, contendo todas as interações entre os actores. Este ficheiro é utilizado para calcular as várias métricas associadas à interacção de actores em fóruns de discussão *online*, bem como para obter o grafo dessas mesmas interações.

4.2.1 Descrição do algoritmo utilizado

De seguida é apresentado o pseudo-código do algoritmo:

Input: matriz com os *posts* ordenados, de tamanho $2 \times N$, onde a primeira coluna são os *ids* dos *posts* e a segunda são os pais dos *posts*;

output: matriz com os níveis, na segunda coluna, e o número de interações


```

interacções ← 0
Para i ← 0 enquanto i < número total de posts
  Se o nível do post actual for diferente de 0 então
    incrementa as interacções em uma unidade
    Para j ← i + 1 enquanto j < número total de posts
      Se o pai to post[j] for post[i] então
        nível do post[j] ← nível do post[i] + 1
Retorna interacções, post

```

A matriz que é passada como *input* ao *script*, é preenchida através da consulta dos *ids* dos *posts* e dos pais destes, na base de dados do Moodle.

Após o retorno da matriz de *input*, passa-se ao preenchimento do ficheiro de *output*, onde vão constar todos os dados da matriz referida. O preenchimento deste ficheiro tem em conta a seguinte condição: caso o *post* actualmente considerado tenha nível *n*, o seu identificador é colocado numa nova linha com *n tabs* a anteceder-lo. Por exemplo, se o nível de um *post* for 0, o seu identificador é colocado no início de uma nova linha, ou seja, com 0 *tabs*. Caso tenha nível 1, o seu identificador é colocado numa nova linha com um *tab* a anteceder-lo.

Após terem sido considerados todos os *posts* incluídos na matriz de *output*, o ficheiro é guardado na pasta *web* do moodle associada aos dados.

4.2.2 Demonstração da execução do *script* de recolha de interacções

4.2.2.1 Apresentação de um fórum de discussão

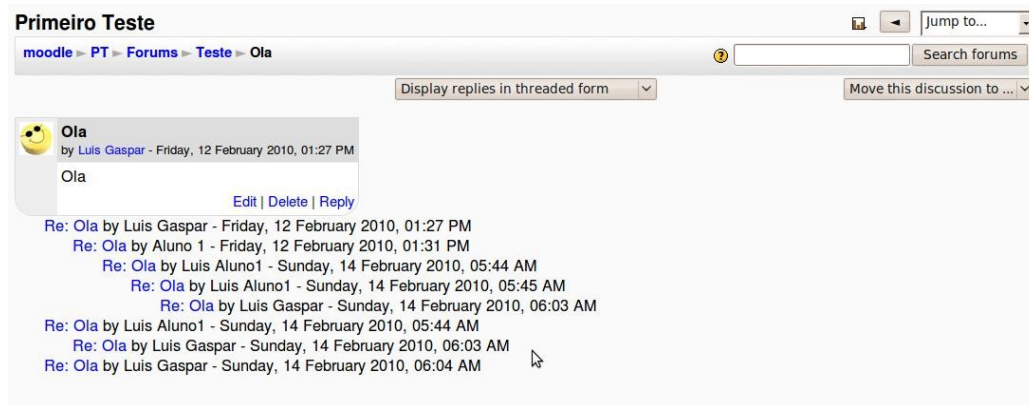


Figura 4.2: Encadeamento de mensagens num fórum de discussão.

A Figura 4.2 mostra as várias interacções no único tópico do fórum de discussão que vai ser utilizado para demonstrar a execução do *script*. Tal como se pode ver nesta figura, o actor Luis Gaspar abre o tópico de discussão, podendo depois perceber-se a série de respostas que se desencadeiam a partir do primeiro *post*.

4.2.2.2 Apresentação do ficheiro de output referente ao fórum anterior

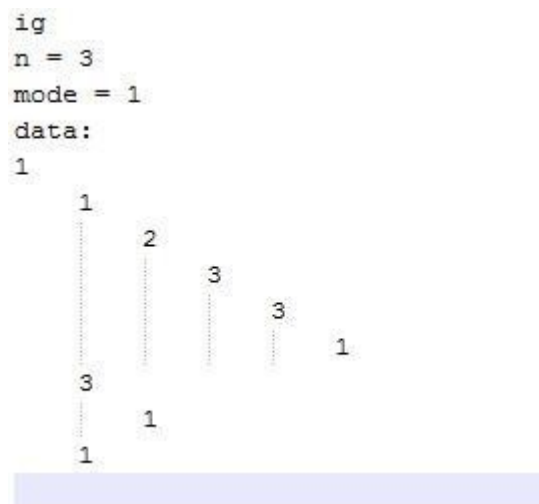


Figura 4.3: Ficheiro de output.

No início do ficheiro (figura 4.3) encontram-se quatro linhas, que indicam, respectivamente:

- O tipo de ficheiro de *output*;
- A indicação do número de actores presentes no fórum;
- O tipo de rede que o ficheiro gerou, nomeadamente, rede de 1-Modo ou rede de 2-Modos;
- Referência ao facto de na linha seguinte começar a descrição das interacções entre os actores.

A seguir a estas, aparecem as linhas que representam as interacções entre actores. Pode ver-se que o encadeamento de *posts* é igual ao apresentado na figura relativa ao fórum exemplo (figura 4.2).

4.3 Implementação do baixo nível

4.3.1 Leitura e armazenamento do ficheiro de interacções

O ficheiro de interacções, obtido através da execução do *script* de recolha de dados num determinado fórum de discussão *online*, tem uma sintaxe bastante simples.

Tal como é referido na secção 4.2.2.2, este contém quatro linhas, antes das interacções, correspondentes a, respectivamente, tipo de ficheiro criado, número de actores do fórum, tipo de rede de representação de interacções e linha de aviso para se perceber que na próxima linha começam as interacções. Destas linhas, vão ser guardados o número de actores e o tipo de rede de representação. Este último dado serve para saber qual é o tipo de representação, por defeito, em que vai ser criado o grafo.

Quanto à parte do ficheiro relativa às interacções entre actores, esta contém os *ids* dos actores que enviam as mensagens. Em primeiro lugar, é feito o mapeamento desses *ids* para caracteres do alfabeto, ou seja, para um grupo de actores com *ids* 1, 2 e 3, passam a estar associados às letras A, B e C. Optámos por este mapeamento, já que, assim, torna-se mais fácil analisar a representação do grafo sem que ocorra uma possível confusão entre o *id* de um actor e o peso de uma determinada ligação, visto serem ambos valores numéricos. Seguidamente, são guardadas todas as interacções, numa lista de listas, sendo que cada lista guarda o número de mensagens enviadas para um determinado actor, bem como os respectivos actores que enviam as mensagens.

O Protótipo

Considerando um determinado actor A, a sua lista tem a seguinte forma:

[A, C, 2, E, 2, F, 2, D, 3, G, 2]

Onde o primeiro elemento indica o actor a que se refere a lista e cada um dos pares de elementos seguintes indicam o actor que envia as mensagens ao actor A e o número de mensagens enviadas para este.

A lista apresentada indica que, para este actor, são enviadas mensagens dos actores C, E, F, D, G. Quanto ao número de mensagens, C, E, F e G enviam 2 mensagens cada um, enquanto que D envia 3 mensagens.

No final da leitura de um determinado ficheiro correspondente às interacções entre actores, a lista de listas guardada tem o seguinte aspecto:

[[C, A, 9, G, 1, F, 2], [A, C, 2, E, 2, F, 2, D, 3, G, 2], [E, E, 1, A, 1], [F, A, 1], [D, A, 1], [G, A, 1], [B]]

Neste exemplo, temos 5 actores que recebem mensagens, nomeadamente, C, A, E, F, D, G e temos o actor B que não recebe mensagens. O actor C recebe mensagens de A (9), de G (1) e de F(2), o actor E recebe 1 mensagem dele próprio e outra mensagem do A e, finalmente, os actores F, D e G recebem uma mensagem do A.

4.3.2 Cálculo do diâmetro dos nós do grafo

Antes de ser calculado o diâmetro dos nós do grafo, correspondente às interacções entre os actores, tem que ser calculado o diâmetro mínimo que pode ser atribuído a um nó que possui interacções, uma vez que, antes o cálculo dos diâmetros de todos os nós, é necessário garantir que o diâmetro mínimo de um nó com interacções é superior ao diâmetro do(s) nó(s) sem interacções, ou seja do(s) nó(s) isolado(s).

Este diâmetro vai ser atribuído ao nó/actor com o menor Índice de Centralização, desde que este índice seja diferente de zero. O cálculo do diâmetro mínimo depende das medidas da *frame* onde vai ser colocado o grafo, nomeadamente, comprimento e altura da *frame*.

A fórmula definida para o cálculo deste diâmetro é a seguinte:

$$D_{min} = comp * alt \div 20000$$

Onde:

- *Dmin* : diâmetro mínimo passível de ser atribuído a um actor;
- *Comp* : comprimento da *frame* onde vai ser apresentado o grafo;
- *Alt* : altura da *frame* onde vai ser apresentado o grafo.

O Protótipo

A divisão por 20.000 da multiplicação entre o comprimento e a altura da *frame*, resulta da experimentação de diversos valores, tendo-se chegado à conclusão que este seria um valor razoável para garantir que, independentemente do tamanho da *frame*, o diversos diâmetros, cujas fórmulas de cálculo são explicadas a seguir, não ficam demasiadamente grandes, tendo em conta o espaço disponível na *frame*.

No que diz respeito ao cálculo dos diâmetros dos nós, este obedece aos seguintes pontos:

1. Atribui-se o valor predefinido ao diâmetro dos nós com IC igual a 0;
2. Ao nó com IC mais baixo, de entre aqueles que têm o IC maior que zero, é atribuído o valor:
 - a. Igual ao diâmetro mínimo, caso este seja superior ao valor predefinido para os nós com IC igual a 0;
 - b. 10% superior ao valor do diâmetro predefinido, no caso contrário.
3. Percorrem-se todos os outros nós, por ordem crescente de IC, e o valor do diâmetro de cada nó é calculado, tendo em conta as seguintes condições:
 - a. Caso o IC do nó for, pelo menos, o dobro do IC do nó anterior:

$$Diam_{act} = (Diam_{ant} * IC_{act} \div IC_{ant}) * 2/3$$

- b. Caso contrário:

$$Diam_{act} = (Diam_{ant} * IC_{act} \div IC_{ant})$$

Onde:

- D_{act} : diâmetro do nó actual;
- D_{ant} : diâmetro do nó anterior;
- IC_{act} : Índice de Centralização do nó actual;
- IC_{ant} : Índice de Centralização do nó anterior.

Ambas as fórmulas apresentadas calculam o diâmetro do nó actual, de modo que a diferença entre os diâmetros actual e anterior seja directamente proporcional à diferença entre o IC actual e o IC anterior. Caso o IC actual seja maior ou igual do que o dobro do IC anterior (primeira fórmula), depois de se efectuar o cálculo anterior, reduz-se esse valor em 1/3.

De referir ainda que este valor, tal como o valor da constante utilizada na divisão no cálculo do diâmetro mínimo, foi alvo de experiências pelo método de tentativa e erro, concluindo-se que este valor seria um valor razoável, tendo em conta que se pretende reduzir o diâmetro a calcular mas sem o reduzir em demasia.

O que se pretende com a utilização de duas fórmulas, em vez de se utilizar apenas uma fórmula, é garantir que os diâmetros dos nós não aumentem demasiado à medida que vão sendo

calculados os diâmetros dos nós com IC maior, por forma a que os diâmetros se mantenham relativamente proporcionais ao tamanho da *frame*.

4.3.3 Representação das interacções no grafo

As linhas, correspondentes às interacções entre actores, podem ligar nós diferentes ou podem também estar direccionadas para o mesmo nó, caso este envie uma resposta a si próprio. Explicamos estas duas situações, que diferem entre si na maneira como essas linhas são desenhadas.

4.3.3.1 Interacções entre actores diferentes

As interacções entre actores diferentes são representadas através de uma linha desenhada entre eles. As coordenadas dos dois pontos, que representam os extremos da linha, são calculadas através das seguintes fórmulas:

$$x = no_x + (D/2 + 1) * \cos(\theta)$$

$$y = no_y + (D/2 + 1) * \sin(\theta)$$

Onde:

- x : coordenada x do ponto que se pretende calcular;
- y : coordenada y do ponto que se pretende calcular;
- No_x : coordenada x do centro do nó em causa;
- No_y : coordenada y do centro do nó em causa;
- D : diâmetro do nó em causa;
- θ : ângulo formado entre a linha que liga os centros dos nós e a horizontal.

Através da utilização destas fórmulas, obtém-se um ponto que se encontra algures em cima da circunferência do nó em causa, variando essa localização consoante o ângulo θ formado entre os dois nós a considerar. Após a obtenção dos dois pontos referentes a cada um dos nós, é traçada a linha entre estes pontos.

Finalmente, verifica-se se os nós em causa têm ligação em ambos os sentidos, ou seja, se ambos os nós enviam mensagens um ao outro. Em caso afirmativo, a linha vai ficar tal como já estava, não lhe sendo acrescentada nenhuma seta em qualquer das extremidades. Caso contrário, é adicionada à linha uma seta na extremidade correspondente ao nó que não envia mensagens ao outro, ficando a seta a apontar para o nó a quem são enviadas mensagens.

4.3.3.2 *Interacções entre o mesmo actor*

Para representar uma interacção entre um actor e ele próprio, é necessário, em primeiro lugar, calcular dois pontos, para, posteriormente, uní-los e formar a curva que representa essa mesma interacção. Esses dois pontos equivalem aos pontos nas extremidades do diâmetro da circunferência do nó, traçado na horizontal.

Após a obtenção desses dois pontos, é traçada uma curva, partindo de um ponto e indo ao encontro do outro ponto, sobrevoando o nó em causa. Para que melhor se perceba a representação deste exemplo de interacção, temos o seguinte exemplo:



Figura 4:4: Representação de uma interacção entre o mesmo actor.

Na figura 4.4, percebe-se facilmente o que foi explicado anteriormente sobre este exemplo de interacção. Como se pode constatar, a figura mostra que a representação da mensagem que o actor E envia para si próprio.

4.3.4 **Definição da espessura das interacções no grafo**

Após a criação da linha que corresponde à interacção, quer seja entre dois nós ou entre o mesmo nó, define-se a espessura da linha, sendo esta proporcional ao número de mensagens trocadas entre os dois actores ou, no caso de ser só entre um actor, é proporcional ao número de mensagens enviadas para ele próprio. A espessura da linha, em *pixels*, equivale ao número de mensagens trocadas entre dois actores ou entre um actor com ele próprio.

4.3.4.1 *Espessura das interacções entre dois actores*

Para garantir que a linha, independentemente da sua espessura, fica centrada com os nós que correspondem aos actores responsáveis por uma interacção, é necessário obedecer aos seguintes pontos:

1. Desenha-se uma linha entre os nós relativos aos actores responsáveis pela interacção, através do método explicado na secção 4.3.3.1;
2. Escolhe-se um desses nós e verifica-se quantas mensagens enviou ao outro nó. Sabendo-se esse valor, desenha-se o número de linhas correspondente ao valor em causa menos um, devido ao facto de já ter sido desenhada a linha do centro e, por isso, uma interacção já estar representada;
3. Partindo da linha inicial, já representada, vão ser desenhadas linhas, alternadamente, acima e abaixo desta (ou à esquerda e à direita se o ângulo for de 90° ou de 270°, ou seja, se a linha que liga os nós estiver na vertical), até atingir o valor anteriormente referido:
 - a. Desenha-se uma linha com a espessura de 1 *pixel* e com 1 *pixel* de deslocamento para cima (ou para a esquerda caso o ângulo for de 90° ou de 270°) em relação à linha inicial;
 - b. Desenha-se outra linha, com igual espessura e igual deslocamento, mas desta vez para baixo (ou para a direita caso o ângulo for de 90° ou de 270°) em relação à linha inicial;
 - c. Incrementa-se uma unidade ao deslocamento e repete-se o processo até o valor ser atingido, guardando-se a posição (acima ou abaixo) onde vai continuar o desenho da interacção, caso o outro nó também tenha enviado mensagens a este, e o deslocamento da última linha desenhada.
4. Escolhe-se o outro nó da interacção e verifica-se quantas mensagens enviou este ao nó anteriormente considerado:
 - a. Caso esse valor seja igual a zero, ou seja, caso não hajam mensagens enviadas desse actor para o outro, conclui-se o desenho da interacção em causa;
 - b. Caso contrário, é iniciado o desenho das interacções:
 - i. Verifica-se em que posição (acima ou abaixo) foi desenhada a última mensagem referente às interacções entre os dois actores em causa, bem como o deslocamento dessa última linha;
 - ii. Desenha-se a primeira linha com a mesma espessura das anteriores (1 *pixel*) e com o deslocamento da última linha desenhada, na posição que foi guardada aquando do final do desenho das linhas referentes ao nó anterior;

O Protótipo

- iii. Repete-se o processo referido aquando do processamento do nó anterior, apenas com a diferença de agora não se subtrair uma unidade ao número de mensagens deste actor para o anterior, uma vez que nenhuma destas mensagens estava representada, ao contrário do que acontecia no caso anterior, sendo já a linha inicial a representação de uma interacção.

Para melhor compreensão do que foi explicado anteriormente, apresentam-se de seguida algumas imagens com a representação de interacções entre dois actores.

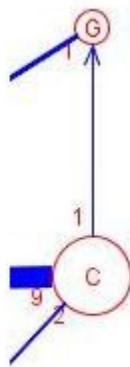


Figura 4:5: Interacções entre os actores C e G.

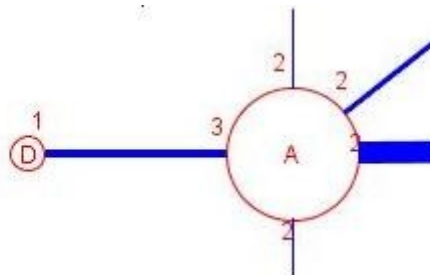


Figura 4:6: Interacções entre os actores A e D.

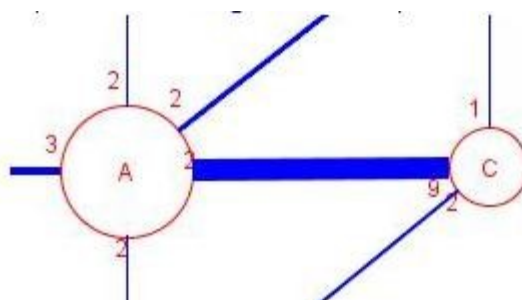


Figura 4.7: Interação entre os actores A e C.

A figura 4.5 apresenta uma interação onde só ocorre o envio de uma mensagem de C para G e, por isso, é apenas desenhada a linha inicial, centrada com os dois nós.

Na figura 4.6 diz respeito a uma interação em que ocorre o envio de 3 mensagens de A para D e uma mensagem de D para A. Sendo assim, a linha que corresponde a estas interações vai ter uma espessura de 4 *pixels*, ou seja, 4 linhas de 1 *pixel*, desenhadas conforme explicado anteriormente.

Quanto à figura 4.7, esta mostra uma ligação entre A e C, sendo que A envia 2 mensagens a C e este envia 9 mensagens a A. Logo, esta interação vai ter uma espessura de 11 *pixels*, ou seja, 11 linhas de 1 *pixel*, desenhadas conforme explicado anteriormente.

4.3.4.2 Espessura das interações entre o mesmo actor

No caso das interações entre um actor e ele próprio, a interação correspondente é representada com uma curva, tal como já foi explicado na secção 4.3.3.2. A espessura desta curva vai ser proporcional ao número de mensagens enviadas do actor para ele próprio, tal como sucede no caso das interações entre dois actores. O cálculo e representação da interação em causa irá ser semelhante ao caso em que estão em causa dois actores diferentes. Tal como no caso anterior, são desenhadas curvas acima e abaixo da curva inicial. É desenhado o número de curvas correspondente ao número de mensagens enviados pelo actor a ele próprio menos um, já que a curva inicial conta como a representação de uma dessas mensagens.

4.3.5 Detecção dos ficheiros de fóruns no directório raiz da aplicação

A detecção de todos os ficheiros “.txt”, referentes a fóruns de discussão e obtidos através da execução do *script* de recolha de dados, é importante, uma vez que, detectando-os em tempo de execução, é possível visualizar qualquer um deles através do desenho do grafo respectivo.

Esta detecção é feita através da análise ao directório raiz do projecto, onde se encontra os ficheiros que o compõem. Percorre-se o directório para encontrar todos os ficheiros que tenham como extensão “.txt”, que é a extensão pré-definida para todos os ficheiros relativos a fóruns de discussão *online*. Os ficheiros com extensão “.txt”, à medida que vão sendo encontrados, serão

guardados num *array* de ficheiros, para que depois possam ser mostrados na interface, em tempo de execução, num dos componentes da interface (explicada na secção 4.4.5).

Por exemplo, se um determinado directório possuir três ficheiros “.txt”, nomeadamente, “grafo.txt”, “grafo2.txt” e “grafo3.txt”, o *output* da função que procura os ficheiros que correspondem a fóruns de discussão é o seguinte:

"Fóruns no directório de raiz: grafo.txt ; grafo2.txt ; grafo3.txt "

4.3.6 Movimentação dos nós no grafo

A aplicação desenvolvida também possibilita a movimentação dos nós, após o desenho do grafo na *frame* correspondente, por forma a que se possa colocar os nós de forma diferente, se assim o pretender. Esta movimentação efectua-se clicando no nó que se pretende mover e, mantendo o clique, arrasta-se esse nó para o local pretendido. Todas as interacções que esse nó tenha com outros actores, ou seja, todas as linhas entre esse nó e outros nós, também vão sofrer com esse deslocamento, acompanhando o nó até à sua nova posição e mantendo a ligação entre os nós intacta.

4.3.7 Algoritmo de desenho do grafo

Uma proposta para o algoritmo de desenho do grafo, correspondente às interacções entre actores num determinado fórum de discussão *online*, que pode ser consultada em Figueira (2009) [FAR09], divide esse desenho em vários pontos, ao longo do seu processamento, nomeadamente:

1. Detecção de cliques;
2. Redução de cliques;
3. Definição do número total de elementos que constituem o grafo;
4. Desenho das órbitas para colocação dos nós e cliques;
5. Desenho dos nós e cliques;
6. Permutação de nós.

4.3.7.1 Detecção de cliques

Os cliques do grafo são detectados através de uma pesquisa exaustiva, percorrendo todos os caminhos possíveis e encontrando, assim, todos os cliques que constam do grafo. Existem algoritmos mais eficientes para efectuar esta pesquisa. No entanto, implementamos desta forma por falta de tempo. Os cliques encontrados são guardados numa lista de listas, contendo várias listas referentes a cada um dos cliques encontrados. De referir ainda que cada uma dessas listas é composta por todos os nós que dizem respeito ao clique que consta nessa lista.

De notar que quando se falam em cliques neste algoritmo é sempre no sentido estrito, já que o sentido lato apenas é considerado na interface da aplicação.

4.3.7.2 Redução de cliques

Partindo da lista de listas com todos os cliques encontrados no ponto anterior, é percorrida essa lista, por forma a encontrar os cliques que, sendo cliques de menor dimensão, estão contidos em algum dos cliques maiores. Esses cliques, depois de encontrados, são eliminados da lista de listas, ficando esta reduzida aos cliques que não estão contidos em mais nenhum dos restantes.

4.3.7.3 Definição do número total de elementos

Neste ponto é definido o número total de elementos constituintes do grafo. Sabendo o número de cliques, através do tamanho da lista de listas obtida anteriormente, e sabendo o número de nós isolados, ou seja, nós que não pertencem a nenhum clique, obtém-se o número total de elementos, através da soma do número de cliques com o número de nós isolados.

4.3.7.4 Definição de órbitas para colocação dos nós e cliques

A definição das órbitas, onde são colocados os elementos (nós e cliques) do grafo, obedece às seguintes condições:

1. Percorrem-se todos os elementos do grafo:
 - a. Se o elemento em causa for um nó, calcula-se o seu Grau de Centralidade, tendo em conta os pesos das suas ligações;
 - b. Se o elemento for um *k-clique*, ou seja, um clique com *k* nós ($k \geq 3$), calcula-se o Grau de Centralidade deste elemento, através da seguinte fórmula:

$$GC = \sum C_i^{GC} / k$$

O Protótipo

Onde:

- GC : Grau de Centralidade, tendo em conta os pesos das ligações;
 - C_i^{GC} : Grau de Centralidade do nó i do clique.
2. Armazemam-se todos os elementos do grafo com o seu Grau de Centralidade respectivo, numa lista de listas. Este armazenamento é feito de duas formas diferentes:
 - a. Se o elemento for um nó, é guardada uma lista que contém este nó e o respectivo Grau de Centralidade;
 - b. Se o elemento for um k -clique, é guardada uma lista com os k nós deste clique e o respectivo Grau de Centralidade do clique.
 3. Efectua-se a ordenação das lista de listas, ordenando-se estas pelo respectivo Grau de Centralidade de cada elemento correspondente a cada uma das listas;
 4. Calcula-se o Índice de Centralização da rede composta por todos os actores;
 5. Calcula-se o diâmetro da primeira órbita (órbita interior), em relação ao centro da *frame*. Este cálculo é efectuado através da seguinte fórmula:

$$D = [0,15 \times comp \times (100 - IC)] \div 100$$

Onde:

- D : Diâmetro da órbita;
 - $Comp$: comprimento da *frame*.
 - IC : Índice de Centralização.
6. Atribui-se ao primeiro elemento da lista de listas com os Graus de Centralidade, ou seja, ao elemento com maior Grau de Centralidade, a órbita definida no ponto anterior;
 7. Percorrem-se os restantes elementos desta lista:
 - a. Caso o elemento actual tiver o mesmo Grau de Centralidade que o anterior, é-lhe atribuída a mesma órbita do anterior;
 - b. Caso contrário, é-lhe atribuída outra órbita, cujo diâmetro é calculado pela seguinte fórmula:

$$D = D_{ant} \times 1,05$$

Onde:

- D : Diâmetro da órbita;
- D_{ant} : Diâmetro da última órbita definida.

8. Guardam-se os elementos do grafo e as suas órbitas atribuídas numa lista de listas. Cada elemento é guardado numa lista cujos elementos são:
 - a. O nós que constituem esse elemento e o diâmetro da órbita correspondente, caso o elemento seja um clique;
 - b. O nó que constitui esse elemento e o diâmetro da órbita correspondente, caso o elemento seja um nó isolado.

De notar ainda que é associado um diâmetro de órbita para cada elemento do grafo, podendo existir elementos com diâmetros iguais, ou seja, elementos com Grau de Centralidade igual. Este facto não leva a uma sobreposição de nós, como é explicado no ponto a seguir.

4.3.7.5 *Desenho dos nós e cliques*

Após a definição das órbitas para o desenho de cada um dos elementos, estes são colocados na sua órbita, tendo em conta os seguintes pontos:

1. Começando pelo primeiro elemento da lista de listas, ou seja, pelo elemento do grafo com a órbita cujo diâmetro é mais reduzido, percorre-se a lista de listas:
 - a. Caso o elemento da lista actual for um *k-clique*:
 - i. Desenha-se o nó, correspondente ao primeiro elemento do clique, colocando o seu centro a coincidir com a órbita, formando um ângulo com o centro da *frame* de 0° em relação à horizontal;
 - ii. Cria-se um *array* para armazenar os ângulos já utilizados, com tamanho correspondente ao número de nós do grafo mais uma unidade, sendo esta unidade relativa ao armazenamento dos ângulos de 0° e de 360° que, por serem iguais, só o ângulo de 0° é considerado, apesar de o de 360° ser utilizado, como se explica a seguir.
 - iii. Guarda-se, no primeiro elemento do *array*, o ângulo de 0° , já considerado, e no último elemento, o ângulo de 360° ;
 - iv. Percorre-se o *array* até se encontrar um elemento vazio e selecciona-se o elemento imediatamente anterior a esse;
 - v. Partindo do elemento vazio, continua-se a percorrer o *array* até ser encontrado um elemento ocupado, seleccionando-se este também;

O Protótipo

- vi. Após a selecção destes dois elementos, calcula-se a média entre eles. Será este o novo ângulo a considerar;
 - vii. Neste caso, como só ainda constam no *array* dois elementos, 0° e 360° , a média entre os dois é de 180° ;
 - viii. Desenha-se outro nó do clique, colocando o seu centro a coincidir com a sua órbita, formando, desta vez, um ângulo de 180° com a horizontal;
 - ix. Guarda-se o ângulo de 180° , numa posição, consoante uma destas hipóteses:
 - (1). Caso o tamanho do *array* seja impar, este ângulo é colocado na posição central;
 - (2). Caso seja par, é calculada a metade do tamanho do *array*, que dá um valor não inteiro, sendo desprezada a parte decimal. Por exemplo, para um *array* de tamanho 13, o ângulo de 180° é guardado na posição 6.
 - x. Selecciona-se o novo ângulo, tendo em conta o que foi referido nos pontos iv, v e vi;
 - xi. Neste caso o ângulo será entre 0° e 180° , logo é de 90° ;
 - xii. Desenha-se o próximo elemento do clique, colocando o seu centro a coincidir com a sua órbita, formando um ângulo de 90° com a horizontal;
 - xiii. Guarda-se este ângulo, considerando as hipóteses do ponto ix;
 - xiv. Repete-se o processo até os k elementos do clique estarem desenhados.
- b. Caso o elemento da lista actual for um nó isolado:
- i. Efectua-se o mesmo que foi referido nos pontos i, ii e iii da alínea a.
2. Percorrem-se os restantes elementos da lista de listas, seguindo as mesmas condições já referidas, quer para o caso de o elemento ser um clique ou um nó isolado. De referir que o *array* não é inicializado quando se termina de desenhar um elemento. Quando se passa para o elemento seguinte, o cálculo do ângulo é igual ao indicado nos pontos iv, v e vi.

4.3.7.6 *Permutação de nós*

O objectivo deste ponto é reduzir, não só a distância entre nós, mas também o número de linhas sobrepostas, tornando assim o grafo mais compreensível. Para isso, consideram-se os cliques que constam do grafo e consoante o número de nós que o constituem, associam-se a estas figuras geométricas, cujo número de vértices é igual ao número dos seus nós. Quer isto dizer que cada um dos nós do clique corresponde a um vértice do polígono.

1. Consideram-se cada um dos cliques do grafo:
 - a. Calcula-se a distância actual de todos os nós do clique para os nós isolados que estão ligados a estes.
 - b. Guarda-se, num *array* com dois elementos, o valor dessa distância e o número das rotações efectuadas. Neste caso, ainda não haviam rotações, logo é colocado o valor 0;
 - c. Efectua-se uma rotação no clique, ou seja, atribui-se cada nó ao vértice seguinte, no sentido horário.
 - d. Após a rotação, calcula-se o novo valor da distância para os nós exteriores:
 - i. Se for menor que a anterior, considera-se esta a melhor posição, sendo substituído o valor, que consta no primeiro elemento do *array*, por este. Altera-se também o segundo elemento do *array* com o número de rotações feitas até se chegar à nova melhor posição;
 - ii. Caso contrário, mantém-se o *array* intacto.
 - e. Repete-se este processo até todos os nós terem sido colocados em todas as posições possíveis do polígono, sendo, no final, considerada, como a melhor posição, aquela que constar no primeiro elemento do *array*.
 - f. Os nós do clique são colocados na sua melhor posição, sendo cada um deles rodado o número de vezes que constar no segundo elemento do *array*.

Com o algoritmo apresentado, cumprem-se todos os princípios enunciados na secção 3.4. Constata-se, por isso, que o algoritmo apresentado, apesar de ser pouco eficiente, cumpre os objectivos inicialmente propostos.

4.3.8 Demonstração da execução do algoritmo

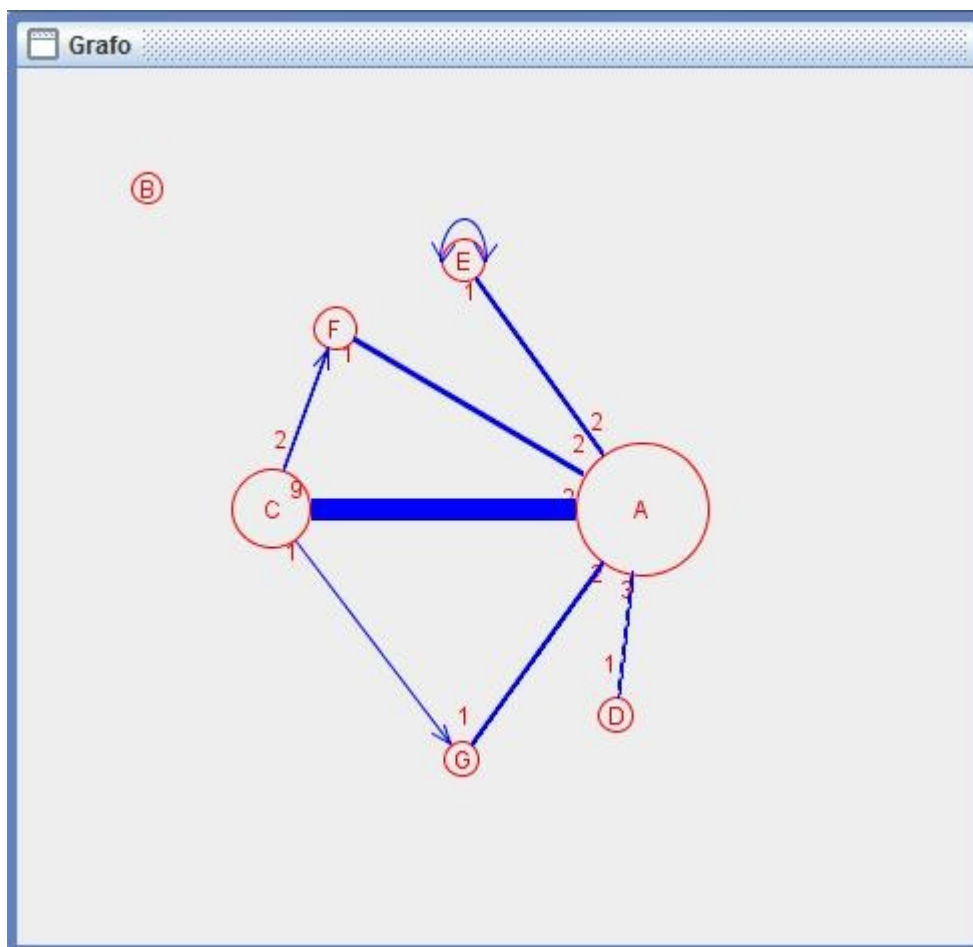


Figura 4:8: Representação de um grafo, utilizando o algoritmo apresentado.

Como se pode constatar na figura 4.8, o actor A, com maior Grau de Centralidade, é desenhado mais próximo do centro da *frame*, formando um ângulo de 0° com a horizontal. O actor C, com o maior Grau de Centralidade, a seguir a A, é desenhado mais afastado do centro do que A, numa posição que forma um ângulo de 180° com a horizontal, sendo este ângulo calculado através das condições referidas anteriormente. Os actores E e F têm o mesmo Grau de Centralidade, sendo, por isso, desenhados na mesma órbita. Vão ser desenhados mais afastados do centro em relação aos anteriores e vão formar um ângulo com a horizontal de, respectivamente, 90° e 135° . Os actores G e D têm também o mesmo Grau de Centralidade. Estes vão ser desenhados na mesma órbita, a uma maior distância do centro, tendo como comparação os restantes já desenhados, e formam um ângulo com a horizontal de, respectivamente, 270° e 315° .

4.4 Visita guiada à interface

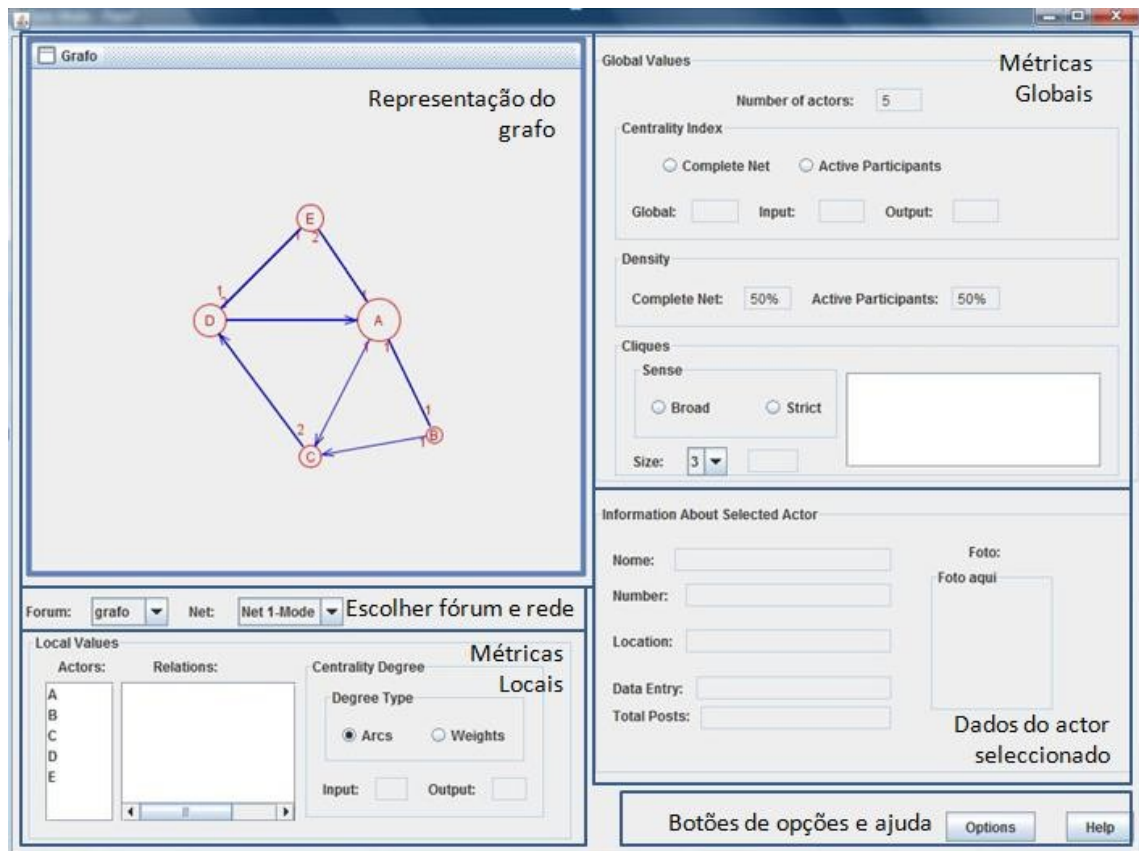


Figura 4.9: Interface da aplicação desenvolvida.

A interface da aplicação, apresentada na figura 4.9 tem os seguintes pontos principais:

- Representação do grafo;
- Métricas locais;
- Métricas globais;
- Informações sobre participantes;
- Escolher fórum;
- Escolher rede;
- Opções disponíveis;
- Ajuda.

4.4.1 Representação do grafo

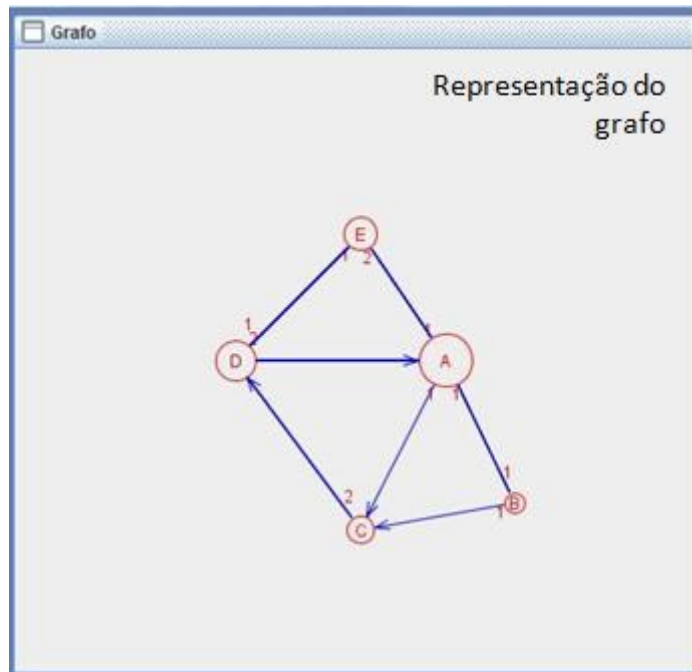


Figura 4:10: Frame de representação do grafo.

Na figura 4.10 consta a secção da interface responsável por apresentar a representação do grafo. Nesta parte da interface é apresentado o grafo correspondente ao ficheiro que foi escolhido através da *comboBox* por baixo do canto inferior esquerdo da *frame* onde vai ser representado o grafo.

Através da representação do grafo, consegue-se perceber o que realmente se passa no fórum, analisando aspectos como: actores centrais ou isolados, cliques e actores que recebem e que enviam mais mensagens.

4.4.2 Métricas Locais

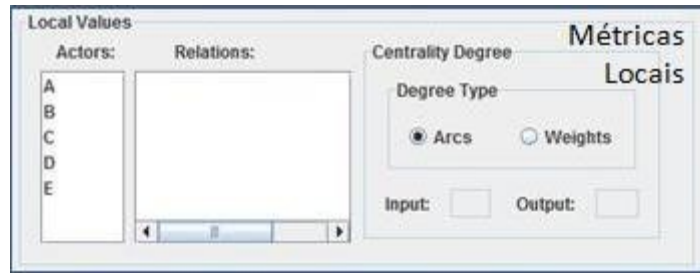


Figura 4:11: Métricas locais.

As métricas locais (figura 4.11), dizem respeito às relações que um determinado nó tem com os restantes, bem como o seu Grau de Centralidade, de entrada e de saída, quer seja contabilizando o número de contactos, quer seja contabilizando o número de mensagens trocadas entre contactos.

Nesta secção da interface, após a escolha de um actor (através de clique num dos actores da lista da esquerda), são apresentadas todas as relações que esse nó tem com os restantes, na área de texto ao lado da lista. Caso o actor não envie nem receba mensagens de alguém, é apresentada, nessa área de texto, uma mensagem identificadora de tal acontecimento.

Do lado direito dessa área de texto, são apresentados os valores relativos ao Grau de Centralidade do actor seleccionado. Caso esteja activa a opção "Arcs", vão ser colocados nos campos de texto, à frente de *input* e *output*, o Grau de Centralidade de entrada e de saída desse nó, respectivamente, tendo em conta apenas o número de contactos. Caso a opção activa seja a opção "Weights" vão ser colocados nos mesmos campos o Grau de Centralidade de entrada e de saída mas, desta vez, tendo em conta o total de mensagens enviadas e recebidas pelo actor, respectivamente.

4.4.3 Métricas Globais

The screenshot shows a software window titled "Global Values" with a subtitle "Métricas Globais". It contains several input fields and radio buttons for configuring network metrics. The "Number of actors" is set to 5. Under "Centrality Index", the "Complete Net" radio button is selected, and the "Global", "Input", and "Output" text boxes are empty. Under "Density", the "Complete Net" and "Active Participants" text boxes are both set to 50%. Under "Cliques", the "Broad" radio button is selected, the "Size" dropdown is set to 3, and there is a large empty text box on the right for displaying results.

Figura 4:12: Métricas globais.

As métricas globais (figura 4.12), estão relacionadas com o número de actores que estão presentes no fórum, com o Índice de Centralização global, de entrada e de saída, tendo em conta toda a rede ou apenas os participantes activos, com a Densidade, tendo igualmente em conta a rede como um todo ou somente os participantes que entram no fórum e, finalmente, com os cliques presentes no fórum, quer seja em sentido lato ou estrito.

Nesta secção, é apresentado o número de actores que se encontram no fórum, a densidade da rede completa e dos participantes activos e, por defeito, os índices de centralidade (global, entrada e saída) da rede completa, podendo, neste ponto, ser escolhida a visualização desses índices tendo em conta os participantes activos. Neste caso, os campos de texto que correspondem a esses índices são preenchidos com os dados relativos aos participantes activos. Quanto aos cliques, escolhe-se se se pretende os cliques em sentido lato ou estrito e escolhe-se também qual o número de nós que os cliques encontrados devem ter para ser apresentados. Posteriormente, os cliques encontrados, que satisfazem as premissas anteriormente definidas, serão apresentados na área de texto à direita dessas opções, sendo também apresentado o número total de cliques encontrados nessas condições, no campo de texto à esquerda da área de texto.

4.4.4 Informação sobre actores



The image shows a web form titled "Information About Selected Actor". It contains five text input fields on the left, each with a label: "Nome:", "Number:", "Location:", "Data Entry:", and "Total Posts:". To the right of these fields is a large rectangular placeholder for a photo, labeled "Foto:" above it and "Foto aqui" inside it. Below the photo placeholder, the text "Dados do actor seleccionado" is displayed.

Figura 4:13: Dados dos actores.

As informações apresentadas sobre o actor dizem respeito ao seu nome, número de aluno, localidade, data em que efectuou o registo no fórum e número total de mensagens colocadas no fórum. A acrescentar a estes dados, é também apresentada uma fotografia do participante, caso esta esteja disponível na base de dados.

Para se ter acesso aos dados de um actor, basta clicar sobre o nó correspondente a esse actor e toda a informação acima descrita vai ser preenchida com os seus dados, através do acesso à base de dados que contem todos os dados dos actores do fórum.

A ligação à base de dados, por forma a permitir a consulta dos dados dos actores para incluir na interface, não foi implementada por falta de tempo.

4.4.5 Escolher fórum



The image shows a web interface with a "Forum:" label and a dropdown menu. The dropdown menu is open, displaying a list of options: "grafo", "grafo2", "grafo3", and "grafo4". The "grafo" option is currently selected and highlighted. To the left of the dropdown, there are labels for "Local Va", "Actor", and a list with "A" and "B".

Figura 4:14: ComboBox para escolha do fórum.

Após a escolha do fórum de discussão que se pretende visualizar, é desenhado o grafo correspondente a esse ficheiro e são preenchidos os dados que aparecem por omissão na interface, nomeadamente, número de actores no fórum, os três índices de centralidade para a rede completa e as densidades para a rede completa e para participantes activos.

4.4.6 Escolher tipo de rede

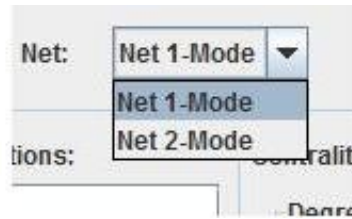


Figura 4:15: ComboBox para escolha da rede.

Nesta secção da interface, é possível escolher se se pretende visualizar a rede de 1-Modo ou a rede de 2-Modos.

Após a escolha da rede que se pretende visualizar, é desenhado o tipo de rede escolhido e são preenchidas as métricas por defeito, já referidas no ponto anterior.

4.4.7 Opções disponíveis

Nesta secção da interface, após o clique no botão "Options", são apresentadas um conjunto de opções de configuração da interface que podem ser alteradas consoante o que se pretenda. Estas opções são:

- Cor dos nós;
- Cor dos arcos;
- Cor de fundo;
- Diâmetros dos nós diferentes ou iguais;
- Espessura das linhas diferentes ou iguais;
- Mostrar ou não mostrar pesos nos arcos;
- Mostrar apenas o nome do nó ou a sequência "nome do nó: número de interações".

Todas estas opções são apresentadas numa janela que abre quando se clica no botão "Options".

4.4.8 Ajuda

Após o clique no botão "Help", irá aparecer uma nova janela contendo todo o tipo de informação relevante para uma correcta compreensão de todos os pontos e funcionalidades de interface. Estas informações podem ser relativas a:

- Representação do grafo;
- Métricas locais;
- Métricas globais;
- Redes de 1-Modo e de 2-Modos.

Com estas informações, pretende-se que, caso ocorram dúvidas sobre os conceitos apresentados na interface, estes conceitos sejam explicados de forma simples e directa, por forma a potenciar a utilização desta interface.

4.5 Applet

A *applet* desenvolvida permite embeber a aplicação em código HTML, tornando, possível a sua utilização na *web*, através da utilização de um *browser*. Para que possa ser possível utilizá-la na *web*, basta importar para a aplicação o ficheiro que se pretende analisar. Para se obter este ficheiro é necessário executar o *script* de recolha de dados, para que este retorne um ficheiro com as interacções do fórum pretendido. De seguida, clica-se no botão criado na *applet* para permitir a importação do ficheiro para a aplicação. Após clicar-se nele, é aberta uma janela onde se pode pesquisar o ficheiro localmente. Finalmente, depois de consumada a importação, a aplicação lê o ficheiro e procede à colocação da representação do grafo e das métricas associadas.

5 Conclusões e Trabalho Futuro

Neste capítulo fazemos um resumo do trabalho desenvolvido, destacando-se todos os aspectos mais relevantes ao longo do seu desenvolvimento. Seguidamente, são apresentadas as contribuições que este trabalho proporciona, nomeadamente, no âmbito de uma análise real dos conceitos da ARS. Após a apresentação das contribuições, segue-se, no pólo oposto, a apresentação das fragilidades que constam no trabalho desenvolvido. Por fim, são apresentados os pontos a melhorar neste trabalho, numa perspectiva de futuro, e que permitam torná-lo cada vez mais importante para o propósito que lhe foi designado, isto é, contribuir para facilitar a análise de interações entre actores em fóruns de discussão *online*.

5.1 Resumo do trabalho

Na revisão do estado da arte, foram focados todos os aspectos que iriam servir de base à elaboração do trabalho. Estes aspectos, na sua grande maioria, estavam contidos numa tese de mestrado já desenvolvida, já que esta era a base do trabalho a ser efectuado. Na realidade, este trabalho pode-se considerar uma extensão dessa tese, já que um dos pontos referidos nela como podendo ser alvo de desenvolvimento futuro é referente aos objectivos propostos por este trabalho. Na revisão do estado da arte, foram também estudadas e justificadas as escolhas das diversas tecnologias para a implementação da aplicação requerida pelo trabalho, apresentando-se todas as escolhas e respectivas justificações. Nesta secção foram também apresentadas as tecnologias mais utilizadas para a ARS, justificando-se o motivo que leva a que não sejam úteis para a realização dos pressupostos deste trabalho.

Na iteração seguinte, foi estudada a forma como são obtidas todas as interações entre actores num determinado fórum de discussão *online*, através da utilização de um *script*, já desenvolvido. Este *script* contabiliza todas as interações entre actores num determinado fórum e as designações dos respectivos actores e coloca-as, ordenadas por níveis, num ficheiro, que, posteriormente, é utilizado na implementação da aplicação pretendida. Todos os pormenores referentes a este *script* encontram-se na secção 4.2.

Sabendo já a sintaxe de retorno do ficheiro obtido com a utilização do *script*, estudou-se as possíveis estruturas de dados para o armazenamento dos actores e respectivas interacções. Chegou-se à conclusão que uma lista de listas seria a melhor opção, sendo fácil o seu preenchimento e a procura no seu conteúdo.

Após a escolha da estrutura de dados para armazenar as interacções entre actores, procedeu-se à implementação do *parser* ao ficheiro. Este *parser* lê do ficheiro as interacções de cada um dos actores e a quem se dirigiram. De referir que, através da leitura do ficheiro são obtidas outras informações, nomeadamente, número de actores envolvidos no fórum de discussão a que corresponde o ficheiro e tipo de rede para representação das interacções (rede de 1-Modo ou rede de 2-Modos).

Tendo concluído a leitura do ficheiro e, por isso, estando todas as interacções entre actores guardadas na estrutura de dados previamente seleccionada, procede-se ao cálculo dos diferentes indicadores numéricos de análise. Estes indicadores dividem-se em três tipos: indicadores de análise local, grupal e global. Dentro da análise local, aparece o Grau de Centralidade que mede o valor de centralidade de um indivíduo tendo por base o número de mensagens em que este está envolvido. Este indicador tem duas variantes, que são o Grau de Centralidade de Entrada e de Saída, variando em relação ao anterior no tipo de mensagens em que se baseia para efectuar o cálculo da centralidade de um indivíduo. O primeiro utilizada para esse cálculo as mensagens recebidas, enquanto que o segundo utiliza as mensagens enviadas. No que concerne à análise grupal, encontram-se o Cliques, indicador que apresenta os grupos que possuem todas as interacções possíveis entre os seus participantes, no sentido estrito, ou que apresenta os grupos em que constam, pelo menos, todas as interacções num sentido entre os participantes desse grupo. Este indicador é utilizado para perceber até que ponto existem grupos de informação num determinado fórum de discussão *online* e, assim, perceber melhor as relações entre indivíduos e eventuais actores isolados no relacionamento entre os participantes do fórum. Quanto à análise global, encontram-se relacionados com esta dois indicadores, que são a Densidade e o Índice de Centralização. A Densidade é calculada tendo por base as interacções existentes e todas as interacções possíveis e serve para analisar se um determinado fórum tem um número de interacções próximo do número de interacções possíveis ou não, ou seja, se a densidade é elevada ou reduzida. O Índice de Centralização é calculado utilizando o Grau de Centralidade de cada um dos participantes no fórum e os Graus de Centralidade dos actores caso esse fórum fosse caracterizado por uma rede em estrela, isto é, uma rede em que um actor está ligado a todos os outros e os restantes só possuem uma ligação com esse actor. Através deste indicador constata-se o quão centralizado está um fórum num único actor. Valores altos deste indicador indicam que a rede está muito centrada num actor e valores reduzidos indicam que o fórum possui vários actores com centralidades semelhantes e que, por isso, a centralidade do fórum se dispersa.

Terminado o cálculo dos indicadores numéricos de análise, procedeu-se à representação das interacções entre actores através do desenho de um grafo. Existem dois tipos de redes para representação dessas interacções, nomeadamente, redes de 1-Modo e redes de 2-Modos. Em relação às redes de 1-Modo, estas representam as interacções entre actores e foram representadas tendo por base estudos psicológicos referentes à facilidade de análise da

representação. O grafo que representa uma rede de 1-Modo é composto por nós que indicam os actores e por arcos que ligam os nós e que representam as interacções entre eles. Para a representação de fóruns em redes de 1-Modo, foi implementado um algoritmo que tem por objectivo representar as interacções entre actores tendo em consideração que este deveria minimizar as sobreposições das interacções, minimizar o espaço entre nós e tornar mais centrais os nós que interagem mais vezes com os restantes. Quanto às redes de 2-Modos, estas são responsáveis pela representação das interacções entre eventos e actores.

Após implementada a representação das interacções nos fóruns de discussão, utilizando os tipos de redes acima descritos, procedeu-se à implementação da *applet* da aplicação, responsável por permitir a visualização da aplicação na *web*. Esta *applet* consiste na aplicação em Java embebida em código HTML, podendo assim visualizar-se a aplicação em qualquer *browser* disponível.

5.2 Contribuições

Este trabalho apresenta os seguintes contributos:

- Criação de uma ligação entre a aplicação de recolha de interacções nos fóruns do Moodle e a interface de visualização das interacções;
- Estruturas de dados para a representação interna das interacções, sob a forma de grafo;
- Representação visual, através de um grafo, das interacções entre participantes de um fórum de discussão *online*;
- Modularidade do grafo, permitindo alterações nas suas propriedades de representação;
- Aplicação que integra a representação visual (grafo) e a representação matemática (indicadores numéricos) para ARS;
- Interface da aplicação sob a forma de *applet*, permitindo a integração numa página HTML.

5.3 Fragilidades detectadas

As fragilidades, entretanto detectadas na nossa aplicação, e passíveis de melhoramento a curto prazo, são as seguintes:

- Algoritmo conducente ao processo de representação dos dados é pouco eficiente. Existem na literatura científica, contudo, vários algoritmos sobre operações em grafos que poderemos incorporar na nossa aplicação e que reduzirão fortemente a complexidade temporal e espacial;
- A ligação da aplicação ao Moodle não está implementada. Na realidade temos as características necessárias para que tal seja feito. Contudo, por manifesta falta de tempo, esse pequeno passo final terá de ficar para trabalho futuro.

5.4 Trabalho Futuro

As possibilidades de melhoramento futuro deste trabalho estão ligadas às fragilidades apontadas no ponto anterior.

Este trabalho pode ser melhorado, tendo em vista a criação do módulo referido na secção anterior, por forma a ser possível a execução da aplicação desenvolvida com este trabalho, integrada na plataforma de ensino Moodle.

Pode também ser melhorado o algoritmo de representação das interacções. Esta melhoria foca-se na eficiência do mesmo, uma vez que chegamos à conclusão que, para fóruns com um número elevado de actores e com uma elevada densidade, o algoritmo implementado torna-se extremamente exigente a nível computacional.

Outra melhoria da aplicação que pode ser feita refere-se à apresentação de dados pessoais de um determinado actor, através da selecção do nó referente a este e posterior consulta à base de dados do Moodle para obter os dados a incluir na interface da aplicação. Este ponto é importante, na medida em que permite uma percepção do actor a que se refere cada um dos nós representados no grafo referente às interacções apresentado na interface da aplicação.

Referências

- [AT09] Analytic Technologies, (2009), Analytic Technologies - Social Network Analysis Software, <http://www.analytictech.com/>
- [BE97] Borgatti, S. P.; & Everett, M. G. (1997). Network Analysis of 2-Mode Data. *Social Networks*, 243-269.
<http://www.analytictech.com/borgatti/papers/borgatti%20%20network%20analysis%20of%202-mode%20data.pdf/>
- [CW09] CriarWeb, (2009), CriarWeb.com Manuais Desenvolvimento web, <http://www.criarweb.com/>
- [ELC08] Cruz, E. L., (2008). *SNSANALYSER: Uma ferramenta para extração e análise de redes sociais a partir de comunidades existentes em sites de relacionamento*. Tese de Mestrado, Universida de Salvador – UNIFACS, Brasil, http://tede.unifacs.br/tde_busca/arquivo.php?codArquivo=311
- [DX09] DevX, (2009), Open Source DataBase Feature Comparison Matrix, <http://www.devx.com/dbzone/Article/29480>
- [FAR09] Figueira, A.R. (2009). *Temporal Online Interactions Using Social Network Analysis*, 4th European Conference on Technology Enhanced Learning. Nice, France.
- [GHW97] Garton, L.; Haythornthwaite, C.; & Wellman, B. (1997). Studying Online Social Networks. *JCMC*, 3(1).
<http://jcmc.indiana.edu/vol3/issue1/garton.html>
- [GV09] Graphviz, (2009), Graph Visualization Software, <http://www.graphviz.org/>
- [JBL08] Laranjeiro, J.B. (2008). *Contributos para a Análise e Caracterização de Interações em Fóruns de Discussão Online*. Tese de Mestrado, Faculdade de Ciências da Universidade do Porto, Portugal, http://www.fc.up.pt/contactos/teses/t_050370107.pdf
- [MIP09] Microsoft, (2009), The Official Microsoft ASP.NET site, <http://www.asp.net/>

Conclusões e Trabalho Futuro

- [MO09] Moodle, (2009), Moodle.org: open-source community-based tools for learning, <http://moodle.org/>
- [NE09] NetBeans, (2009), Welcome to NetBeans, <http://netbeans.org/>
- [PHP09] PHP, (2009), PHP: Hypertext Preprocessor, <http://php.net/index.php>
- [QA05] Quiroga, A. (2005). Introducción al Análisis de Datos Reticulares: Prácticas com UCINET6 y NetDraw1, <http://revista-redes.rediris.es/webredes/talleres/redes.htm>.
- [SLS09] SlideShare, (2009), Upload & Share PowerPoint presentations and documents, <http://www.slideshare.net/>
- [STG92] Streeter, C. L.; & Gillespie, D. F. (1992). Social Network Analysis. *Journal of Social Service Research*, 16(1/2), 201-222. <http://www.utexas.edu/courses/streeter/fall2005sw393t19/readings/sna.doc>
- [SUN01] Sun, (2001), Java 2 Platform, Standard Edition, v 1.3.1 API Specification, <http://java.sun.com/j2se/1.3/docs/api/>
- [SUN09] Sun, (2009), JavaServer Pages Technology, <http://java.sun.com/products/jsp/>
- [WIK09] Wikipedia, (2009), the free encyclopedia, <http://en.wikipedia.org/>